



Innovative assessment across the disciplines

An analytical review of the literature

FINAL REPORT

Dai Hounsell, Nancy Falchikov, Jenny Hounsell,

Monika Klampfleitner (University of Edinburgh)

Mark Huxham, Karen Thomson (Napier University)

Sheena Blair (Glasgow Caledonian University)

Notes on contributors

Dr Sheena Blair is Head of the Division of Occupational Therapy, School of Health and Social Sciences, Glasgow Caledonian University. Within the Innovative Assessment project she was particularly responsible for the theme 'portfolios and other unconventional writing assignments' and for the literature on Health Sciences and Practices.

Dr Nancy Falchikov is an Honorary Fellow in the Department of Higher and Community Education, Moray House School of Education, University of Edinburgh. Within the Innovative Assessment project she was particularly responsible for the theme 'student involvement in assessment' and for the literature on Geography, Earth and Environmental Sciences.

Dr Dai Hounsell is Professor of Higher Education in the Centre for Teaching, Learning and Assessment, Moray House School of Education, University of Edinburgh. He designed, led and coordinated the Innovative Assessment project and also took responsibility for the theme 'what feedback and guidance is given, by what means, and when' and for the literature on Business, Management, Accountancy and Finance.

Jenny Hounsell is a Research Officer in the Centre for Teaching, Learning and Assessment, Moray House School of Education, University of Edinburgh. Within the Innovative Assessment project she was particularly responsible for the theme 'assessment of groups and collaboration' and the literature on History, Classics and Archaeology.

Dr Mark Huxham is a Senior Lecturer and Senior Teaching Fellow in the School of Life Sciences, Napier University. Within the Innovative Assessment project he was particularly responsible for the theme 'presentations and other non-written assignments' and for the literature on Bioscience.

Monika Klampfleitner was a Research Associate in the Centre for Teaching, Learning and Assessment, Moray House School of Education, University of Edinburgh, working on the Innovative Assessment project, creating and managing the database and literature searches, and with responsibility for the literature on Languages, Linguistics and Area Studies. In April 2007 she took up a new position as Advisor (Education) to the Vice-Chancellor, Chemnitz University of Technology.

Dr Karen Thomson is a Lecturer in Higher Education, Educational Development, Napier University. Within the Innovative Assessment project she was particularly responsible for the theme 'the use of new technology in assessment' and for the literature on Psychology.

Contents

	Executive summary	4
	i Introduction	4
	ii Delimiting the scope and focus of the search-and-review	4
	iii The creation of the database	5
	iv Analysis by themes	6
	v Analysis of selected subjects	9
	vi Conclusions, implications and recommendations	10
1	Introduction	13
2	Delimiting the scope and focus of the search-and-review	15
	Defining 'innovative' assessment	16
	What's meant by 'UK' literature?	17
	Classifying by subject area	17
	Devising a typology of themes	18
	The genre dimension	18
	Fine-tuning the review design	24
3	The creation of the database	24
	EndNote field: genre	27
	EndNote field: form of innovation (category)	27
	EndNote field: focus of innovation	27
	EndNote field: subject grouping	28
	EndNote field: context	29
	EndNote field: level/year of study	30
	EndNote field: mode of assessment	30
4	Analysis by themes	30
	4.1 Introduction	30
	4.2 Presentations and other non-written assessments	31
	4.3 Portfolios and other unconventional writing assignments	35
	4.4 Assessment of groups and collaboration	38
	4.5 Student involvement in assessment	42
	4.6 Use of new technology in assessment	51
	4.7 What guidance and feedback is given, by what means, and when	54
5	Analysis of selected subjects	58
	5.1 Introduction	58
	5.2 Bioscience	58
	5.3 Business, Management, Accountancy and Finance	59
	5.4 Geography, Earth and Environmental Sciences	60
	5.5 Health Sciences and Practice	61
	5.6 History, Classics and Archaeology	62
	5.7 Languages, Linguistics and Area Studies	64
	5.8 Psychology	65
6	Conclusions, implications and recommendations	66
	Innovation by theme	68
	Theme drivers	68
	Thematic points of reference	69
	Theme implications	69
	Innovation by subject area	70
	Innovation by genre and fitness-for-purposes	70
	Accounts of practice	71
	Empirical studies	72
	Enhancement projects and evaluations	73
	End-uses of the literature	73
	Review impact and sustainability	73
7	Acknowledgements	75
8	References	75

Executive summary

i Introduction

In universities, where the choice of methods of assessment is devolved, such methods can differ by subject area, by course and by institution. Yet while devolution offers scope to reconfigure assessment approaches, there may be correspondingly greater difficulties in keeping track of where and how changes in assessment practices are taking place. The growing volume of publications on developments in assessment offers one important means of keeping abreast, but how such publications are used and valued depends to a substantial extent on one's professional standpoint. Practitioners may prize relevance to practice and the significance of the topic or issue studied, while researchers may be more concerned with the research process rather than its product.

Against this backdrop, the aims of this analytical review were:

- to map documented innovations by type of innovation, by discipline/subject area, and by *genre* or form of publication, while being alert to the variety of potential users of the literature and the end-uses to which it might be put
- to collate and evaluate, within and across the types of innovation differentiated, the conceptual perspectives and frameworks utilised, the nature of the evidence deployed, and wider issues and concerns raised.

The review cannot, however, make an unqualified claim to comprehensiveness, given that its timescale placed constraints on the range of subjects and years of publication covered.

ii Delimiting the scope and focus of the search-and-review

Defining what is meant by *innovative* assessment is notoriously tricky since its meaning varies from one subject area or institution to another and over time. The present review followed an inclusive and pragmatic approach: innovative assessment was taken to be that which was novel in the eyes of its begetters or beholders; entailed more than a minor modification; was not self-evidently traditional; and had been documented over the last decade.

The focus was on publications with UK authorship, and/or on innovative assessment in UK higher education settings. Within the literature, priority was given to mainstream, peer-reviewed sources (chiefly journal articles), but accommodating many *greyer* publications where they met minimum reporting requirements, were accessible, and had a non-trivial contribution to make to the literature.

The search was built around disciplinary clusters as represented in the Academy's Subject Centres and encompassed the full span of subject areas, except Medicine. Since the review was to focus on innovative assessment *across* the disciplines, publications that traversed subject areas were given higher priority.

A grounded typology of innovative forms and directions comprising twelve principal categories was developed, and a novel attempt was also made to differentiate publications by type of publication or *genre*, on the grounds that the criteria that might be used to evaluate the merits of a publication falling within one genre would

not necessarily be appropriate to appraising the quality of one from a different genre. The approach to characterising genres drew on the work of Swales (1990) and Greenhalgh (2001). Ten genres were identified, defined chiefly by their main communicative purposes:

<i>Genre</i>	<i>N</i>
<i>Account of Practice.</i> A publication that aims to describe and reflect on an instance of a change or development in day-to-day professional practice in assessment, and is usually self-reported and self-evaluated by one or more subject practitioners.	141
<i>Commentary/Opinion Piece (including a contribution to debate).</i> A publication that aims to argue for a reappraisal of, or a significant change to, one or more aspects of contemporary assessment practices, processes or policies.	4
<i>Compendium (of evolving practices).</i> A publication that seeks to chart, document and illustrate contemporary and evolving assessment practices and procedures, within or across subject areas, institutions or sectors.	7
<i>Empirical Study.</i> A publication that aims to report the findings of a research enquiry, investigation, experiment or survey of assessment practices, processes or policies.	74
<i>Enhancement Project.</i> A publication aiming to report the outcomes of a joint initiative by subject practitioners and higher education researchers/academic developers to monitor and improve assessment practices or processes.	19
<i>Evaluation.</i> A publication that seeks to report the outcomes of an independent (third-party) data-gathering evaluation of an assessment initiative or aspect of assessment practice or policy.	7
<i>Guide to Professional Practice (including textbooks).</i> A publication that attempts to provide a general overview of and orientation to assessment purposes, principles, practices, processes, procedures and/or policies, or to an aspect of one or more of these.	9
<i>Guidelines.</i> A publication that seeks to offer practitioners and others [including administrators, academic managers, external examiners and quality reviewers] guiding principles, precepts or rules-of-thumb on assessment.	13
<i>Review of the Literature.</i> A publication that seeks to appraise and synthesise the findings and outcomes of other publications on one or more aspects of assessment, usually to arrive at an up-to-date picture of what is known and understood about the topic.	14
<i>Theory/Conceptualisation.</i> A publication that seeks to contribute to theoretical advance, whether by proposing a new conceptualisation of assessment processes, practices or policies, or by modifying or refuting an existing one.	51
Total <i>N.B. Where a publication falls into more than one category it has been counted under each of the categories concerned</i>	339

iii The creation of the database

In its completed form, the Innovative Assessment EndNote database comprises 317 references to publications on innovative assessment. Each reference provides key bibliographic information and reports genre, context, form and direction of the reported innovation from 1996 onwards. A comprehensive search-and-review of articles on innovative assessment was carried out, with a particular focus upon key

cross-disciplinary higher education journals. These represent two thirds of all references in the database, with the remaining one third consisting of book chapters and subject-specific publications.

iv Analysis by themes

In conducting the search-and-review, the goal was not only to catalogue individually each instance of innovative assessment, but also to chart its main innovative forms, directions or trends, and the relative incidence of each based on the numbers of publications retrieved. The former could be achieved through item-by-item indexing by keyword. The latter entailed building an overall map or typology:

<i>Form of innovation (Category)</i>	<i>N</i>
1. Innovation in assessment generally	1
1.1 assessment practices and processes generally	6
1.2 assessment regimes	9
1.3 perceptions and experiences of staff, students and others	8
1.4 enhancing assessment practices	5
2. Modes of assessment and balance between them	5
3. Conditions under which the assessment is carried out, including when and where	5
4. The nature of the task assessed including:	
4.1 presentations and other non-written assessments	27
4.2 portfolios and other unconventional writing assignments	49
4.3 assessment of groups and collaboration	37
4.4 assessment of performance	0
4.5 assessment of students on placements/in the workplace/in situ	11
4.6 assessment of projects and dissertations	1
5. Criteria used in the assessment and their relative weighting	29
6. Student involvement in assessment	58
7. Use of new technology in assessment	45
8. What guidance and feedback is given, by what means, and when	40
9. How assessment tasks and processes are organised, aligned and managed	31
10. Assessment policies and codes of practice	4
11. Regulatory frameworks and procedures	14
12. How students' performance and achievements are combined, aggregated, recorded and represented	16
Total	
<i>N.B. Where a publication falls into more than one category it has been counted under each of the categories concerned</i>	401

Of the 12 main categories of innovative themes analysed, eight themes spread across six categories accounted for nearly four-fifths of the total number of relevant publications retrieved and recorded in the database: student involvement in assessment; portfolios and similar assignments; the use of new technology in

assessment; the provision of guidance and feedback; assessment of groups and collaboration; how assessment tasks are organised, aligned and managed; criteria and their relative weighting; and presentations and other non-written assignments. A noteworthy but unanticipated topic which also emerged from the search-and-review was that of refashioning assessment to engage with greater student diversity, including the needs of non-traditional UK students, students with disabilities or students from overseas.

Presentations and other non-written assessments

Most items within this theme concerned student presentations, mainly by individual students. There were relatively more articles from the arts and humanities than from science. All studies cited the need for teaching communication skills as a motivation. A common theme was the notion of a *public performance* being particularly motivating (as well as scary) for students, and the idea of openness and interactivity in questioning leading to critical understanding. A number of the studies discussed student involvement in the negotiation and setting of assessment criteria, as well as the marking of the presentations. The importance of clear criteria was raised in many of the papers, partly as simply good practice, but also reflecting wariness over the possible subjectivity and unreliability of oral versus written assessment.

Portfolios and other unconventional writing assignments

This theme included publications on portfolios, learning journals, reflective or reflexive journals, projects, diaries, log books and different forms of textual practices. The two largest subject groupings were vocational: Education, and Health Sciences and Practice. Assessment of portfolios could be problematic, and the necessity for clear criteria-referenced frameworks was endorsed by several papers. There were also logistical issues such as agreement on how to structure portfolios, confidentiality issues and how to manage the more personalised aspects of learning, how to give feedback timeously and how to achieve inter-rater reliability. Current resource issues within higher education made it challenging to deal with such large quantities of written material, despite the perceived benefits which were felt to include: documentation of learning over time; a resource for future learning; improving written communication skills; and subscribing to the values of lifelong learning.

Assessment of groups and collaboration

Reported reasons for innovation on this theme were: helping students to develop skills needed in their transition from school to university: introducing a work-related element to group work: experimenting with assigning students to groups and the consequent effect on grades: and developing on-line group collaboration. Few of the innovations extended beyond one course or department, and examples from the social sciences predominated. Authors were mostly positive about their own and their students' perceptions of the usefulness of assessing groups and collaboration in developing a range of skills, including the ability to cooperate as a team, to think critically and to reflect on their learning. Practical issues raised included how to allocate marks to groups and whether it was the product, the process or both which should be assessed. Departmental and institutional policies would have been needed in order to phase and blend group assessment appropriately.

Student involvement in assessment

A wide variety of disciplines were using and researching ways of involving students in assessment, most frequently through peer assessment. The most frequently occurring advice offered by authors concerned: the need for preparation and training for both students and teachers; making clear the rationale for involving students in assessment; and following an incremental approach or using pilot studies. A few authors deemed their implementation time-consuming, but some expressed the belief that the time spent was worth it. Some authors identified the need for further study, more development and evaluation, while others called for staff support, for support materials or for the need to create a supportive atmosphere to encourage students to air their concerns. Some authors claim their implementation to be beneficial in aiding the development of personal and lifelong learning skills. Other benefits claimed include helping students overcome unrealistic expectations, helping shared understanding and encouraging integration.

Use of new technology in assessment

More than half of the innovations reviewed under this theme were concerned with multiple-choice assessment and its variants. Discussion boards, the management of institutional risk in the context of electronic assessment, the submission of portfolios presented electronically and the submission of assessment using multimedia were also considered. A wide variety of subject areas were represented, although many innovations had a general focus. Many of the implications suggested by the authors reflect established good practice in assessment more generally, including: alignment of assessment to learning outcomes; clarity of purpose of assessment; and the importance of clear assessment criteria. Other implications specifically for electronic assessment were: the need for adequate resources to establish the assessment; the need for training of both staff and students; and the management of the ensuing dialogue about learning.

What guidance and feedback is given, by what means, and when

About two-thirds of the items reviewed were concerned with the provision of feedback, with varying emphases on: students' perceptions and experiences; their use of feedback; effects on performance; enhancing feedback; students' grasp of tutors' expectations; and assessment criteria. Some items focused on guidance in the context of supervision, while others were concerned with particular strategies for providing, communicating or enhancing feedback, including the use of computer-based spreadsheet feedback, multiple-choice questions, model answers, one-minute papers in lectures, evolving tutorial files and forms of technology-supported feedback. The largest subject grouping was that of Business, Management, Accountancy and Finance. Two key implications were the need for much greater recognition of the challenges in providing constructive and effective feedback, and the desirability of greater attention to tutor-student dialogue and interchange about expectations.

v Analysis of selected subjects

The items retrieved were also categorised by subject area, chosen to match those of the Higher Education Academy Subject Centres:

	<i>N</i>
Art, Design and Media	6
Bioscience	19
Built Environment	3
Business, Management, Accountancy and Finance	38
Dance, Drama and Music	5
Economics	1
Education	32
Engineering	14
English	14
Geography, Earth and Environmental Sciences	22
Health Sciences and Practice	27
History, Classics and Archaeology	10
Hospitality, Leisure, Sport and Tourism	3
Information and Computer Sciences	22
Languages, Linguistics and Area Studies	5
Law	3
Materials	0
Maths, Statistics and Operational Research	11
Philosophical and Religious Studies	4
Physical Sciences	2
Psychology	14
Social Policy and Social Work	9
Sociology, Anthropology and Politics	8
Total	
<i>N.B. Where a publication falls into more than one category it has been counted under each of the categories concerned</i>	272

The best represented subjects were: Business, Management, Accountancy and Finance; Education; Health Sciences and Practice; Geography, Earth and Environmental Sciences; and Information and Computer Sciences. Medicine was excluded because the literature on innovative assessment within Medicine was considered to be quite extensive, as well as more accessible (at least to those within that field) than was likely to be the case in many other subject areas. Of the remaining 23 subject areas, the literature from seven were analysed in more depth: Bioscience; Business, Management, Accountancy and Finance; Geography, Earth and Environmental Sciences; Health Sciences and Practice; History, Classics and Archaeology; Languages, Linguistics and Area Studies; and Psychology.

The analyses showed that each of the subject areas could be classified by a range of the themes discussed above, with no one theme being particularly common in any of the disciplines. Likewise, innovations in assessment took a range of forms in each of the subject areas and, although based in one area, could largely be applicable across all the disciplines. While many of the journal articles reviewed were from higher education journals, there were also articles from subject-specific journals, as well as book chapters and conference proceedings.

The Academy Subject Centre websites provided an additional source of literature on assessment, although these varied in the amount and type of information available. Some have their own electronic journal covering assessment as well as other aspects of teaching and learning, and some carry website links to relevant sources of literature on assessment. Most of the websites also contain some guidelines, case studies and/or briefing papers on the topic of innovations in assessment.

vi Conclusions, implications and recommendations

The results of this review indicate clearly that the UK literature on innovative assessment is large, buoyant, richly varied and diffuse. The publications retrieved spanned the subject range, focused on a wide array of themes, represented a diversity of genres and communicative purposes, and were spread across publication sources. To seasoned observers of assessment trends, the most prominent themes within the literature may come as little surprise. A less anticipated theme, however, was the emergent topic of refashioning assessment to engage with greater student diversity, by addressing the needs of UK students from non-traditional backgrounds, students with disabilities, or students from overseas countries, particularly where English was not their first language.

Drivers or stimuli to innovation in assessment varied widely, within and across themes. Recurring stimuli included a desire to recraft assessment practices the better to reflect contemporary mass higher education with its large and diverse cohorts of students and pressures on resources; an interest in capitalising upon developments in information technology or methods of assessment; a heightened emphasis on developing transferable skills needed in the graduate workplace and for lifelong learning; and a more general commitment to sustaining and enhancing the quality of students' learning, which was sometimes linked to a particular pedagogical philosophy.

Points of reference adduced by the authors were also richly varied, drawing on *how to* guides to teaching and assessment in higher education; the subject literature or similar accounts of practice by others; conceptual and empirical studies of assessment and student learning; or wider theoretical perspectives on pedagogy including (amongst others) constructivism, self-regulation, identity and academic literacies. There was also substantial variation in the degree to which authors were knowledgeable about relevant sources, whether conceptual, empirical or practical. In some instances references were simply sparse; more frequently what had been cited seemed fortuitous rather than judiciously selected. While this is not surprising, given that so many of the authors were from a subject area other than education, it has the effect of working against a truly cumulative literature and the evolution of a shared understanding of what is known and understood within the field. The present review and database should help to enhance familiarity with the literature, but might have a greater impact if allied by

the Academy to briefing seminars for prospective author-innovators, subject advisers and academic developers.

The practical, procedural and policy implications identified by the authors of the publications retrieved also differed widely in ways that were closely bound up with the theme concerned and with practices in the particular course settings where innovations had taken place. However, three wider implications stood out: the necessity for careful groundwork, briefing and training of both students and staff if initiatives were to take firm root; the need to attend to other salient aspects of course organisation and management, including time and resources; and changes needed at the departmental or institutional level if some innovations in assessment were to flourish in other courses and programmes of study.

Analysis of documented innovations by subject area showed that the best-represented subject areas were those of business and management, education and health sciences and practice, while seven subject areas were each represented by four or fewer publications. There were no obvious patterns of incidence related to relative age of a discipline, vocational versus non-vocational fields, or to broader subject groupings. How far documentation of initiatives mirrors the scale of innovation in assessment within a given subject area is open to question, but the incidence of documentation within a subject area may reflect how vigorously innovative initiatives are promulgated and debated, while their relative absence might act as a brake on potential innovation, by furnishing few exemplars from which others might learn. There may therefore be merit in the adoption by the Academy of a capacity-building strategy, which brokered partnerships between subject areas with relative low and relatively high levels of documented innovation.

A guiding principle of this review has been to recognise the extent to which the literature of innovative assessment not only has a profusion of origins but a diversity of destinations, needs and end-uses. A novel genre-based approach to the review was therefore adopted to avoid appraising the literature wholly or principally on research-driven criteria. Accounts of practice were the most common genre (141 items). Empirical studies were also strongly represented (74 items), as were theoretical or conceptual contributions (51 items). There were also 19 enhancement projects, 14 reviews of the literature and 13 guidelines. Commentaries, compendia, evaluations and guides to professional practice were each represented by fewer than ten publications.

The accounts of practice analysed could have potentially substantial shortcomings: they were typically self-reported and self-evaluative, and there appeared to be no well-understood conventions or guiding principles for them which authors, referees or editors might rely on. Yet as reports by insiders of authentic course-based experiences, they played a vital role in disseminating within and beyond a subject what has been learnt from developments in practices, and in inspiring others to emulate them. Four key features characterised the best accounts of practice: the authors made explicit their own involvement in the initiatives being related; the accounts were located within the relevant literature; distinctive features of the settings for innovations were highlighted; and the scope and robustness of the data underpinning the accounts were made explicit. In the light of these observations, the Academy might wish to consider how the overall quality of accounts of practice and their potential value to practitioners could best be enhanced. Possible avenues included the drafting of a code of practice, ideally

in consultation with relevant editors and publishers, producing guidance materials, and organising a rolling programme of seminars and workshops.

Empirical studies were sometimes hard to distinguish from accounts of practice that had gathered data from beyond the immediate. However, sample sizes were for the most part small and restricted, as were the range of settings typically surveyed; and while the better examples of empirical enquiry were alert to their limitations, there were also less satisfactory examples prone towards incautious generalisation. Here too there may be scope for the Academy to enhance the genre, e.g. through measures to build research capacity amongst subject teachers and academic developers, and to try to *scale up* sample sizes by fostering alliances between aspiring researchers. Neither strategy, however, may suffice to address a fundamental weakness of research in this field: the virtual absence of follow-up studies that track the impact of assessment approaches on students' progression over successive years or levels of study.

Both the review and the database can play an important role not only in mapping documentation of innovations in assessment in higher education, but also in contributing to capacity-building. The database should enable those pursuing and documenting innovative initiatives to familiarise themselves more readily with what has been attempted elsewhere and what lessons can be drawn, within and across subject areas. At the same time, the genre-based analyses should help alert end-users of the literature to the communicative and evidential strengths and limitations of publications from different genres. It is also hoped that the typology of genres developed for this particular review may prove to be more widely applicable.

There is, nonetheless, a pressing need for systematic empirical investigation of how the pedagogical literature of higher education is accessed, made sense of and put to use for practical and policy-related purposes – a cluster of issues that take on very particular significance for the Academy given its role as a gatekeeper and intermediary between the worlds of practice, policy and research. Happily, there are a small number of examples from higher and other sectors of education where this has been attempted, and which could assist in exploring possible ways of proceeding.

1 Introduction

- 1.1 A distinctive and much-prized characteristic of higher education is that the choice, not only of curriculum content and teaching-learning strategies, but also of methods of assessment is to a considerable extent devolved. This in turn means that assessment methods differ not simply by subject area (as is also the case in secondary and further education): they also vary from university to university, and – even in the same department or faculty – may take different forms from one module, course unit or degree programme to another.
- 1.2 An important consequence of the devolution of assessment (and, it can reasonably be claimed, a substantial benefit of it) is that there is potentially much larger scope to fine-tune, overhaul or reconfigure assessment approaches, whether in response to the push of constrained resources or new policy emphases, the pull of developments in information technology, or the outcome of rethinking within a course team about what could and should be learnt by the students following a particular programme of study. However, a greater and highly localised capacity to innovate can also increase considerably the difficulties of keeping track of where and how changes in assessment practices are taking place, when viewed from a range of standpoints or stakeholder positions. The latter range widely, including for instance:
- the university teacher attracted by a novel approach to assessment and keen to find out about, and learn from, the experiences of colleagues elsewhere who have already experimented with it
 - the new head of department who would like to ascertain how well the department compares with others in the same discipline in the variety of assessment strategies deployed and how these are blended
 - the subject adviser who is committed to keeping abreast of developments in assessment both within that discipline and in cognate subject areas, in order to offer up-to-date and well-informed guidance to others
 - the members of a curriculum validation and review board who, faced with proposals for an approach to assessment within a course that represents a significant departure from current practice, want to be confident that potential limitations as well as advantages have been anticipated and steps taken to safeguard academic and professional standards
 - the researcher or evaluator with a particular interest in tracking, mapping or systematically investigating innovative approaches to assessment, whether within a given discipline or in a cross-section of subject areas.
- 1.3 In principle, and regardless of one's particular professional standpoint, the burgeoning pedagogical literature of higher education might seem to offer a fertile and relatively direct means, not only of finding out about the what and where of innovative assessment initiatives, but also of examining what might be learnt from these. In practice, however, things are far from straightforward. The literature of higher education, as of education generally, is not only abundant but typically diverse and disparate, with a multiplicity of providers and channels of scholarly

and professional communication (Hounsell, 1987; see also Tight, 2003, 2004). It can therefore be relatively inaccessible to anyone who does not have ready access to specialist library holdings of scholarly journals, monographs and conference proceedings, or who is not adept at navigating their way through a constellation of academic, professional and institutional websites. A further confounding factor is that the literature takes many different and potentially confusing forms: it ranges from reports of empirical investigations by formally trained educational researchers, through reflections by subject teachers on their professional practice in a given course setting, to guidelines on policies and practices by quality assurance or professional accreditation bodies.

- 1.4 Furthermore, how these myriad forms of publications are viewed – and even more crucially, how they are *valued* – also depends to a substantial extent on one's professional standpoint and the kind of use to which one intends to put the literature. In traditional educational research and scholarship, it is the systematic empirical study which has tended to be prized, and through it the pursuit of generalisation across individuals, levels of study, subject areas, settings and even cultures. Conversely, and from this same perspective, practice-focused reports and reflections by teachers and other practitioners are likely to be judged of much lower merit, or even disregarded, on the grounds that they offer evidence that is anecdotal, they lack academic rigour, and they are restricted to very specific settings. To the mainstream subject teacher, however, this *hierarchy of evidence* (Greenhalgh, 2001) can appear lopsided or inappropriate. Here generalisation may be a lesser priority than contingency – the precise particulars of the context within which the work reported took place and their salience for its outcomes – just as the trappings of research and scholarship may count for rather less than the *street credibility* afforded by a fellow practitioner with expertise in that subject area and a depth of experience of teaching and assessing students at the same or a similar level.
- 1.5 Adrianna Kezar (2000) makes a similar observation: in her study of U.S. researchers' and practitioners' perspectives on the higher education literature, practitioners tended to prize relevance to practice and the significance of the topic or issue studied, while researchers were more concerned with the research process rather than its product. In other words, the two groups of stakeholders had different criteria for what made the literature significant – a finding which takes on even greater significance when seen against the broader shift underway in the research-practice nexus, where the focus is not only on the researcher-as-disseminator, but also on the practitioner-as-learner (Nutley, Walter and Davies, 2003).
- 1.6 It was against this challenging backcloth that the present review was undertaken. Commissioned by the Higher Education Academy following competitive tendering, it sought to map and critically appraise documented innovations in assessment in higher education, as reported in the UK literature, and spanning the disciplinary spectrum. More specifically, the aims of the analytical review were:
 - to map the incidence and distribution of documented innovations by type of innovation, by discipline/subject area, and by genre or form of publication, while being alert to the variety of potential users of the literature and the end-uses to which it might be put

- in the same vein, to collate and evaluate, within and across the types of innovation differentiated, the conceptual perspectives and frameworks utilised, the nature of the evidential warrant claimed, and issues and concerns raised likely to be of wider significance for assessment policies, procedures and practices.

- 1.7 The review was undertaken by a cross-institutional, cross-disciplinary team drawn from Edinburgh, Napier and Glasgow Caledonian Universities. From the outset, two principal outcomes of the review were specified. First, an EndNote database would be compiled comprising key bibliographic information on all the publications identified in the course of the review, and subsequently made widely available by the Academy. Second, the present report would be drafted, reporting the results of analyses of the literature retrieved by innovative theme, discipline and genre. Further and subsequent dissemination of the review findings would take the form of conference and seminar presentations and journal articles.
- 1.8 The timescale stipulated by the Academy for the completion of the review was the period 1 September 2006 to 31 March 2007. This inevitably placed constraints on what could feasibly be accomplished within that time-span, including an early decision to exclude Medicine from the subject areas covered on the grounds that the literature on innovative assessment within Medicine was considered to be quite extensive, as well as more accessible (at least to those within that field) than was likely to be the case in many other subject areas. A further constraining factor was the sheer volume of UK publications on innovative assessment in the other subject areas that were retrieved in the course of the review. That in turn meant taking difficult decisions while the review was in train, to ensure that it would be completed on schedule. As a consequence, the review cannot make an unqualified claim to comprehensiveness, in spite of the large volume of publications it has identified, catalogued and analysed.
- 1.9 Note should also be taken of the combined effect of a brisk timescale and the extent to which a review of this kind and on this theme entailed a journey into comparatively uncharted territory. Consequently (as will become clearer in the sections which follow), various difficult decisions about definitions and boundaries had to be provisionally settled at the outset but then refined *on the hoof*, evolving in the light of the publications retrieved and reviewed.
- 1.10 The remainder of this report is in five parts. Following fuller accounts of how the review was designed and conducted, and of how the database was created, analyses of the publications retrieved are given for a selection of innovative themes and for a cross-section of subject areas. The report closes with conclusions, implications and recommendations.

2 **Delimiting the scope and focus of the search-and-review**

- 2.1 This section of the report highlights and examines key factors in determining the scope and focus of the search-and-review strategy adopted. These factors

included adopting a working definition of *innovative assessment*; clarifying what would constitute a UK publication (a dual task); choosing an approach to classifying by subject area; devising a tentative typology of themes; and tackling the crucial but elusive dimension of publication genre. The section concludes with a discussion of how the review design was refined in the light of these factors and the already noted constraints of time and resources.

Defining 'innovative' assessment

- 2.2 Assessment in higher education is widely perceived to have been undergoing a sea-change over the last 10-15 years, as has frequently been recorded (see for example, Brown and Glasner, 1999; Gibbs, 1995; Hounsell, McCulloch and Scott, 1996; Bryan and Clegg, 2006). Yet precisely what is meant by *innovative* assessment, or *alternative* assessment as it is sometimes called (e.g. Struyven et al. 2003; Maclellan, 2004) is notoriously tricky. A distinction is sometimes made between *radical* innovation and *routine* or incremental change (see for example McKenzie et al., 2005), but applying such distinctions to a richly diverse higher education system is fraught with difficulties. What is considered innovative for one subject area or institution may not be so for another (Hounsell and McCulloch, 1999), and the degree of innovativeness of any initiative is inescapably transitory, fading with the passing of time. As Hannan and Silver have observed:

Innovation in higher education has generally been taken to mean a planned or deliberate process of introducing change, directed towards (but not necessarily achieving) improvements or solving or alleviating some perceived problem. Such changes may be new to a person, course department, institution or higher education as a whole. An innovation in one situation may be something already established elsewhere, but its importance for this discussion is that initiative takers and participants see it as an innovation *in their circumstances*.

(Hannan and Silver, 2000, p. 10)

- 2.3 Southwell and her colleagues note similar varieties of interpretation, and adopt the crisp and clear definition that an innovation is "an idea, product, process or service that adds value, and is useful or transforms current practice in the context to which it is applied" (Southwell et al., 2005, p. 17).
- 2.4 The present review takes a similarly inclusive and pragmatic approach. Broadly speaking, that which is innovative in relation to assessment is taken to be that which is novel in the eyes of its begetters or beholders, and entails more than a minor or trivial adjustment or modification. However, the beholders for the purposes of this review are taken to mean the UK higher education sector as a whole rather than any one institutional or professional group or subject field, and that necessitates avoiding too narrowly based a definition in embarking on a search of the literature and in attempting to categorise innovative developments in assessment thematically. As a second guiding principle, therefore, our search-and-review has focused on any assessment forms, approaches or practices which are not self-evidently traditional, where *traditional* encompasses the conventional assessment diet of assigned, in-course essays and reports, and unseen, written end-of-course examinations (Hounsell, Xu and Tai, 2007).

- 2.5 Thirdly, given that innovation is transient, the search-and-review has concentrated on publications published in the course of the preceding decade, i.e. from 1996 onwards. The one exception is in the thematic analyses (section 4 below), where reference is made as and when appropriate to influential conceptual and empirical contributions that may have been published prior to 1996.
- 2.6 Fourthly, as far as empirical and conceptual studies of assessment are concerned, the approach taken has been to include rather than exclude them from the database in borderline cases (e.g. studies of student or staff perceptions and experiences of assessment where the focus is on, say, assessment experiences in-the-round rather than solely of innovative assessment practices). This can also be justified on the grounds that any work which advances knowledge and understanding of assessment practices is *ipso facto* innovatory.

What's meant by 'UK' literature?

- 2.7 Our brief was to focus on the UK literature. We have taken UK authorship and/or a focus on innovative assessment in UK higher education settings as the definitive criteria. Priority has also been given (as described more fully in section 3. below) to material published in the UK, with the exception of the abundant journal literature, as UK-based higher education journals are not confined to articles of UK provenance, just as international ones may contain material by UK authors.
- 2.8 What constitutes a *publication* or *the literature* in this field has also called for careful consideration, since material on innovative assessment can be found in profusion not only in traditional, blue-riband, scholarly sources such as journals and monographs, but also in a growing *grey literature* of paper and electronic publications produced by a host of professional organisations and associations, institutional teaching-learning and academic development centres, and R & D projects and programmes. To exclude all grey literature would be to overlook a potentially rich source of insights into – and a key influence upon – changes in assessment practices. On the other hand, to have attempted to cover it with any degree of comprehensiveness would not have been practicable within the time-frame and resources at our disposal. The approach chosen was to give priority to mainstream, peer-reviewed sources (most especially, to journal articles), but accommodating many greyer publications where these had opportunistically come to our attention, and with the proviso that they met minimum requirements for reporting practices, were publicly accessible, and had a non-trivial contribution to make to the understanding of evolving assessment practices in UK higher education.

Classifying by subject area

- 2.9 In delimiting disciplines, a course has to be steered between dysfunctionally over-clustering subjects and unmanageably wide differentiation. The approach adopted here was an avowedly pragmatic one built around disciplinary clusters as represented in the Academy's Subject Centres. This had the merit of being a well-understood typology amongst practitioners as well as researchers, while facilitating consultation with Subject Centres about potential resources within their

particular fields. For reasons noted earlier, the search encompassed the full span of subject areas except Medicine.

- 2.10 In addition to searching for, and recording in the database, relevant publications from the resulting 23 subject areas, it was felt to be worthwhile to undertake a fuller analysis of publications within a representative range of subject areas. The outcomes of these analyses are presented in section 5 below. These analyses also make reference to further subject-specific publications available through the Subject Centres concerned. No attempt has been made to record all of these in the Innovative Assessment database, nor to explore subject-specific publications in a wider range of disciplines, since the search-and-review – as its title makes clear – is chiefly concerned with innovative assessment *across* rather than within disciplines.

Devising a typology of themes

- 2.11 In conducting the search-and-review, the goal was not only to catalogue individually each instance of innovative assessment, but also to chart its main innovative forms, directions or trends, and the relative incidence of each. The former could be achieved through item-by-item indexing by keyword. The latter entailed building an overall map or typology. This was initially devised using seven main categories as a starting-point: the mode of assessment; when and where assessments took place; the nature of the task assessed; who was involved in assessing; how new technology was used to facilitate assessment; how assessment tasks and processes were organised and managed; and assessment policies, codes of practice and mechanisms for monitoring and review.
- 2.12 This tentative categorisation was then extended and refined recursively, as the search-and-review progressed, resulting in the typology shown in Table 2.1.

The genre dimension

- 2.13 Intrinsic to this analytical review is an attempt to differentiate publications by type of publication or genre, on the grounds that the criteria that might be used to evaluate the merits of a publication falling within one genre (e.g. an empirical study or a theoretical work) would not necessarily be appropriate to appraising the quality of one from a different genre (e.g. a review of the literature or a textbook). A genre-based approach also has the merits of potentially greater inclusivity in the types of publications covered (i.e. by not focusing primarily or solely on the research-oriented literature), and of making it much easier for potential users of the resulting database to home in on those types of publications which best match their needs and preferences.

Table 2.1: Typology of forms of innovative assessment

	<i>Themes selected for in-depth analysis</i>
1. Innovation in assessment generally	
• assessment practices and processes generally/as a phenomenon	
• assessment regimes	
• perceptions and experiences of staff, students and others	
• enhancing assessment practices	
2. Mode of assessment (e.g. diagnostic/formative/summative/ipsative) and balance between them	
3. Conditions under which the assessment is carried out, including when and where	
4. The nature of the task assessed including:	
• presentations and other non-written assessments	X
• portfolios and other unconventional writing assignments (e.g. logs, blogs, diaries)	X
• assessment of groups and collaboration	X
• assessment of performance (dance, art, drama, music)	
• assessment of students on placements/in the workplace/in situ	
• assessment of projects and dissertations	
5. Criteria used in the assessment and their relative weighting	
6. Student involvement in assessment	X
7. Use of new technology in assessment	X
8. What guidance and feedback is given, by what means, and when	X
9. How assessment tasks and processes are organised, aligned and managed (e.g. to tailor assessment to students' backgrounds, needs and aspirations; to foster progression; to match methods to curriculum goals, to ensure consistency of marking)	
10. Assessment policies and codes of practice	
11. Regulatory frameworks and procedures (including plagiarism, reviews, appeals)	
12. How students' performance and achievements are combined, aggregated, recorded and represented (including PDP, progress files and degree classification)	

- 2.14 The genres identified for the purposes of this review are outlined and illustrated in Tables 2.2 and 2.3, and derive their definition from the work of Swales (1990). As conceptualised by him, what characterises a particular genre is a shared set of communicative purposes and conventions:

A genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognised by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. [...] In addition to purposes, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience.

(Swales, 1990, p. 58).

- 2.15 Thus the typical communicative purpose of an empirical study, for example, would be to report the findings of a research enquiry, experiment or survey, while a review of the literature would aim to appraise and synthesise the findings and outcomes of other publications to arrive at an up-to-date understanding of the topic.
- 2.16 From the standpoint of conventions, parallel differences (such as focus, structure, approach and conceptual underpinning) could be anticipated in what was expected of authors seeking to publish work in either genre, and – concomitantly – what was expected by readers when engaging with work within the genre concerned. For some well-established genres – including the two examples just given – these conventions are by and large widely known and understood (even if they are contested and refashioned from time to time). However, there can be other genres where the ground-rules are by no means clear or widely accepted. In the case of the literature on innovative assessment, one genre in particular needs to be highlighted because, although generally overlooked, it is quite commonplace: the account of practice.
- 2.17 An account of practice, it is suggested, is a publication that aims to report and reflect on an instance of day-to-day professional practice – as more narrowly defined here, one that involves a significant change or development in assessment practice in a particular subject and institutional setting. It represents a challenging genre because it usually involves academics stepping out of a familiar and well-established role for which they have undergone a long apprenticeship (communicating their research and scholarship within their chosen discipline), to a much more uncertain one for which they have may well have had little or no preparation (communicating their experiences, insights and reflections on their practice beyond as well as within their discipline). Accounts of practice can also present difficulties because they are quintessentially self-mediated: the authors of the account are also the originators of the initiative they are reporting and reflecting upon, and the reader is therefore wholly reliant on whatever information or evidence they choose to present, as well as how they interpret this. Contrast this with reports by *outsiders* rather than *insiders*, as in the genres of the evaluation, the empirical study, the review of the literature — or indeed the compendium of changing practices, which typically draws on a range of accounts of practice in surveying and illustrating innovative developments on a specific theme.

Table 2.2: Publications on innovative assessment in higher education: a typology of genres

<p>ACCOUNT OF PRACTICE</p> <p><i>A publication that aims to describe and reflect on an instance of a change or development in day-to-day professional practice in assessment, and is usually self-reported and self-evaluated by one or more subject practitioners.</i></p>
<p>COMMENTARY/OPINION PIECE (including a contribution to debate)</p> <p><i>A publication that aims to argue for a reappraisal of, or a significant change to, one or more aspects of contemporary assessment practices, processes or policies.</i></p>
<p>COMPENDIUM (OF EVOLVING PRACTICES)</p> <p><i>A publication that seeks to chart, document and illustrate contemporary and evolving assessment practices and procedures, within or across subject areas, institutions or sectors.</i></p>
<p>EMPIRICAL STUDY</p> <p><i>A publication that aims to report the findings of a research enquiry, investigation, experiment or survey of assessment practices, processes or policies.</i></p>
<p>ENHANCEMENT PROJECT</p> <p><i>A publication aiming to report the outcomes of a joint initiative by subject practitioners and higher education researchers/academic developers to monitor and improve assessment practices or processes.</i></p>
<p>EVALUATION</p> <p><i>A publication that seeks to report the outcomes of an independent (third-party) data-gathering evaluation of an assessment initiative or aspect of assessment practice or policy.</i></p>
<p>GUIDE TO PROFESSIONAL PRACTICE (including textbooks)</p> <p><i>A publication that attempts to provide a general overview of and orientation to assessment purposes, principles, practices, processes, procedures and/or policies, or to an aspect of one or more of these.</i></p>
<p>GUIDELINES</p> <p><i>A publication that seeks to offer practitioners and others [including administrators, academic managers, external examiners and quality reviewers] guiding principles, precepts or rules-of-thumb on assessment.</i></p>
<p>REVIEW OF THE LITERATURE</p> <p><i>A publication that seeks to appraise and synthesise the findings and outcomes of other publications on one or more aspects of assessment, usually to arrive at an up-to-date picture of what is known and understood about the topic.</i></p>
<p>THEORY/CONCEPTUALISATION</p> <p><i>A publication that seeks to contribute to theoretical advance, whether by proposing a new conceptualisation of assessment processes, practices or policies, or by modifying or refuting an existing one.</i></p>

Table 2.3: A typology of genres: some examples from the assessment literature

<p>ACCOUNT OF PRACTICE McCreery, C. (2005) 'Less is more: rethinking assessment in a first-year history unit.' <i>Synergy</i> (Sydney Univ) no. 22. http://www.itl.usyd.edu.au/synergy/article.cfm?articleID=265</p>
<p>COMMENTARY/OPINION PIECE (including a contribution to debate) Elton, L. (2004) 'Should classification of the UK honours degree have a future?' <i>Assessment & Evaluation in Higher Education</i>, 29(4), pp.415-422</p>
<p>COMPENDIUM (OF EVOLVING PRACTICES) Juwah, C. et al., (2004) <i>Enhancing Student Learning through Effective Formative Feedback</i>. (SENLEF Project). York: Higher Education Academy, 28-30. http://www.heacademy.ac.uk/senlef.htm</p>
<p>EMPIRICAL STUDY Hyland, P. (2000) 'Learning from feedback on assessment', in Booth, A. and Hyland, P. (eds.) <i>The Practice of University History Teaching</i>. Manchester: Manchester University Press, pp. 233-247.</p>
<p>ENHANCEMENT PROJECT Rust, C. Price, M. & O'Donovan, B. (2003) 'Improving students' learning by developing their understanding of assessment criteria and processes.' <i>Assessment & Evaluation in Higher Education</i>. 28(2), 147-164.</p>
<p>EVALUATION Hall, R. (2002) 'Aligning learning, teaching and assessment using the web: an evaluation of pedagogic approaches'. <i>British Journal of Educational Technology</i>. 33(2), pp. 149-158</p>
<p>GUIDE TO PROFESSIONAL PRACTICE (including textbooks) Brown, G., Bull, J. and Pendlebury, M. (1997) <i>Assessing Student Learning in Higher Education</i>. London: Routledge.</p>
<p>GUIDELINES QAA (2006) <i>Code of Practice for the Assurance of Academic Quality and Students in Higher Education. Section 6: Assessment of students</i>. Gloucester: Quality Assurance Agency. Nicol, D. and Macfarlane-Dick, D. (2006) 'Formative assessment and self-regulated learning: a model and seven principles of good feedback practice'. <i>Studies in Higher Education</i>. 31.2, pp. 199-218.</p>
<p>REVIEW OF THE LITERATURE Falchikov, N. and Goldfinch, J. (2000) 'Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks.' <i>Review of Educational Research</i>. 70.3, pp. 287-322.</p>
<p>THEORY/CONCEPTUALISATION Sadler, D.R. (1989) 'Formative assessment and the design of instructional systems'. <i>Instructional Science</i>. 18, pp. 119-144.</p>

2.18 However, limitations can also be strengths, depending on the purposes and values of a diverse readership. Thus the insider account may be little-valued by a research-minded reader looking for the kind of third-party, systematic gathering and synthesis of data to be found in an archetypal empirical study. However, to fellow-insiders – other professional practitioners working in the same or a cognate subject area, for example – it can be precisely this aspect of such accounts that make them credible and authentic. They are by those working 'at the chalk-face' of assessment — those who, as Ecclestone and Swann (1999) put it, "possess intimate day-to-day contextual knowledge, including an understanding of those aspects of the situation which are not easily measurable, or even readily observable" (pp. 387-388).

Table 2.4: Reading accounts of practice-in-context – key issues

<i>Aspect</i>	<i>Key Issue</i>	<i>Scope of Variation</i>
PROPRIETY	<i>What is the nature of the stake held in the account (and in the practice to which it relates) by those who have reported it?</i>	Self-reported, self-evaluated, versus independently reported and independently evaluated.
INCIDENCE	<i>To what extent is the account presented particularised (e.g. linked to a specific individual or group, in a specific locale or institution, at a specific point in time)?</i>	Focus on a single reference-group (e.g. of learners) task, activity, locale and/or subject area versus focus on multiple reference-groups, tasks, activities, locales and/or subject areas.
CONTEXTUALISATION	<i>To what extent does the account presented seek to identify and depict distinctive and localised features of the setting within which the practice took place?</i>	Little or no contextualisation or 'scene-setting' versus extensive, detailed contextualisation.
GROUNDING IN THE LITERATURE	<i>How far is the account set within a review of what is already known and understood about the practice documented?</i>	Little or no scholarly grounding versus partial or patchy grounding versus extensive and systematic grounding.
CONCEPTUALISATION	<i>What attempt is made to deploy salient existing conceptual frameworks or to generate new ones?</i>	Considered only relating to issues of practice or policy versus informed by relevant concepts or theories.
DATA GENERATION AND COLLECTION	<i>What attempts were made to pinpoint, gather and analyse systematic data concerning the practice reported?</i>	Reliance on own experience, observations or insights and incidental or opportunistic evidence versus purposive, proactive data-gathering.
CRITICAL ENGAGEMENT	<i>How far does the account seek to problematise, analyse and critically evaluate the practice, and/or its reporting and analysis of it?</i>	Narrative and descriptive versus reflective, analytical and critical.

Source: Dai Hounsell, University of Edinburgh, EdD Programme, April 2007 [revised]

- 2.19 What set of criteria, then, might be appropriate to reviewing the communicative fitness-for-purpose of an account of practice, in the absence of well-established conventions for that genre? As an initial baseline, the analysis by genre will draw on a framework of questions developed by one of the authors (DH) for use within a professional a doctorate in education in appraising accounts of practice. The framework is shown in Table 2.4.

Fine-tuning the review design

- 2.20 The principal aims of the analytical review have already been outlined in the Introduction above (section 1.6). These initial aims formed the basis for a set of more keenly focused research questions (RQs), further refined in the opening phase of the project, and set out in Table 2.5. As the table also shows, the first group of questions (RQ1-RQ3) were to be applied to the database as a whole. The second group of research questions (RQ4-RQ7) were addressed to the items retrieved under each of the themes selected for more in-depth analysis.

3 The creation of the database

- 3.1 The Innovative Assessment EndNote database comprises 317 references of publications on innovative assessment. Each reference provides key bibliographic information and addresses research questions 1-3 by recording genre, context, form and direction of the reported innovation. These annotations allow a refined literature search and should make the database a valuable tool for the wide range of users who are concerned with assessment in higher education. This section of the report describes what is included in the database, and explains and analyses the information it provides.
- 3.2 The database documents the UK literature, as defined in section 2, from 1996 onwards. Influential literature published prior to 1996 and non-UK literature is therefore not included in the database.
- 3.3 The setting of these boundaries allowed for a comprehensive search-and-review of articles on innovative assessment in the five cross-disciplinary, core higher educational journals:
- Active Learning in Higher Education
 - Assessment & Evaluation in Higher Education
 - Higher Education
 - Studies in Higher Education
 - Teaching in Higher Education

Table 2.5: Research questions

<i>Question no.</i>	<i>Topic</i>	<i>Research question in full</i>	
RQ.1	Forms and directions of innovation in assessment across the disciplines	How can we best characterise and classify the differing forms and directions that documented innovation in assessment across the disciplines has taken in higher education over the last decade?	Address question to whole database
RQ.2	Distribution of forms/directions of innovative assessment	How are these forms and directions of documented innovation in assessment distributed – across disciplines, by levels/years of study, and over time?	
RQ.3	Documentation of forms/directions of innovative assessment	How are these forms and directions of innovation in assessment in the disciplines documented in the literature – not only by mode of publication, but by genre?	
RQ.4	Innovative drivers and stimuli	For given forms/kinds of documented innovation in assessment, what seem to be the main drivers of, or stimuli to, innovation? [e.g. policy drivers, resource pressures, concerns about quality, desire to find a way of assessing a specific quality or skill; the pull of a technological development]	Address question to selective themes
RQ.5	Conceptual, empirical and practical points of reference	For given forms/kinds of documented innovation in assessment, what seem to be the main conceptual/empirical/practical points of reference adduced? [In other words, what have the innovations' originators/authors claimed to have taken as guides or compass bearings to help them see or navigate their way forward?]	
RQ.6	Implications for assessment practices, procedures and policies	For given forms/kinds of innovation in assessment, what seem to be the main implications (including main helps and hindrances) identified by the authors/originators?	
RQ.7	Observations and recommendations on documenting of innovative assessment	Reviewing the material retrieved overall, what observations and reflections can be made on its robustness and fitness-for-purpose (in supporting and advancing assessment practices, policies and research), and how might its value be enhanced?	Address question to selective themes and whole database

- 3.4 A total of 208 articles from these educational journals are included in the database. This represents two-thirds of all references. The remaining one-third (109 items) consists of other educational literature, e.g. book chapters and subject-specific publications. The latter was searched and reviewed, in particular, regarding the main themes which are analysed in section 4.
- 3.5 Three-fifths of the included articles from the five core higher education journals are published in *Assessment & Evaluation in Higher Education*, making it not only the most important source within the educational journals, but with almost 40% of all references also the single most important source for the database. Table 3.1 shows the distribution of the articles across the five educational journals.

Table 3.1: References in higher educational journals

<i>Sources</i>	<i>N</i>	<i>%</i>
Active Learning in Higher Education	21	10.1
Assessment & Evaluation in Higher Education	126	60.6
Higher Education	2	1.0
Studies in Higher Education	22	10.6
Teaching in Higher Education	37	17.8
Total educational journals	208	100.0

- 3.6 Given our search strategy, it is not surprising that journal articles represent 82% of all references. The table below (Table 3.2) presents the distribution across reference types within the database.

Table 3.2: References by type of item

<i>Reference Type</i>	<i>N</i>	<i>%</i>
Book	6	1.9
Book section	25	7.9
Case study	1	0.3
Conference proceedings	2	0.6
Journal article	260	82.0
Other/generic	22	6.9
Report	1	0.3
Total database items	317	100.0

- 3.7 In addition to the bibliographic information including keywords and an abstract, each reference provides information on genre, context, form and direction of the reported innovation (RQs 1-3). Where information on the context is not reported in a publication, the corresponding EndNote field is blank. The EndNote fields which address research questions 1-3 are discussed below.

EndNote field: genre

- 3.8 The references are categorised according to the genres outlined in section 2. Accounts of practice are by far the most common way of reporting innovation. They constitute 42% of all references, followed by empirical studies with 22% (Table 3.3). The literature does not always fit neatly into those categories and some hybrids are assigned two genres. While this reflects the complexity of the literature, it complicates the statistical analysis. It means that n=339 for genre is higher than the number of references on the database (n=317). This strategy of assigning more than one category where appropriate is also applied to other EndNote fields where a publication retrieved falls into more than one of the relevant categories. The percentages in the following tables are therefore percentages of occurrences rather than of the number of publications.

Table 3.3: References by genre

<i>Genre</i>	<i>N</i>	<i>%</i>
Account of practice	141	42.6
Commentary/opinion piece	4	1.2
Compendium (of evolving practices)	7	2.1
Empirical study	74	21.8
Enhancement project	19	5.6
Evaluation	7	2.1
Guide to professional practice	9	2.7
Guidelines	13	3.8
Policy document	0	0.0
Review of the literature	14	4.1
Theory/conceptualisation	51	15.0
Total	339	100
<i>N.B. Where a publication falls into more than one category it has been counted under each of the categories concerned</i>		

EndNote field: form of innovation (category)

- 3.9 This field indicates the form and direction of the innovation reported (Table 3.4). The literature is classified according to the typology of themes presented in section 2. If the innovation documented is tangent to two or three themes, all are listed, the main theme coming first in the list. A piece of literature was always reviewed by the two reviewers responsible for the main theme. Therefore, the analyses of selected themes in section 4 are based on publications that focus on these topics.

EndNote field: focus of innovation

- 3.10 The innovative aspect of a publication with regards to assessment is summarised in one sentence to inform the user of the database at a glance.

Table 3.4: References by form of innovation

<i>Form of innovation (category)</i>	<i>N</i>	<i>%</i>
1. Innovation in assessment generally	1	0.2
1.1 assessment practices and processes generally	6	1.5
1.2 assessment regimes	9	2.2
1.3 perceptions and experiences of staff, students and others	8	2.0
1.4 enhancing assessment practices	5	1.2
2. Modes of assessment and balance between them	5	1.2
3. Conditions under which the assessment is carried out, including when and where	5	1.2
4. The nature of the task assessed including:		
4.1 presentations and other non-written assessments	27	6.7
4.2 portfolios and other unconventional writing assignments	49	12.2
4.3 assessment of groups and collaboration	37	9.2
4.4 assessment of performance	0	0.0
4.5 assessment of students on placements/in the workplace/in situ	11	2.7
4.6 assessment of projects and dissertations	1	0.2
5. Criteria used in the assessment and their relative weighting	29	7.2
6. Student involvement in assessment	58	14.5
7. Use of new technology in assessment	45	11.2
8. What guidance and feedback is given, by what means, and when	40	10.0
9. How assessment tasks and processes are organised, aligned and managed	31	7.7
10. Assessment policies and codes of practice	4	1.0
11. Regulatory frameworks and procedures	14	3.5
12. How students' performance and achievements are combined, aggregated, recorded and represented	16	4.0
Total		
<i>N.B. Where a publication falls into more than one category it has been counted under each of the categories concerned</i>	401	100.0

EndNote field: subject grouping

- 3.11 The disciplinary clusters as represented in the Academy's Subject Centres are used to map the distribution of innovations across the disciplines. About two-thirds (215) of the 317 database references are linked to one or more subject areas. Of those, Business, Management, Accountancy and Finance, and Education are the subject areas which are best represented in the core higher education journals. The distribution of subject areas differentiated by education journals and other sources is shown in Table 3.5. Where publications refer to disciplines in general

terms, e.g. humanities, sciences, no subject grouping is allocated, but the information is noted in the EndNote field: context.

Table 3.5: References by subject area and source

	N (includes some double-counting)	% key journals (n=273)	% other sources (n=273)	% of all subject grouping entries(n=273)
Art, Design and Media	6	1.5	0.7	2.2
Bioscience	19	5.1	1.8	7.0
Built Environment	3	0.7	0.4	1.1
Business, Management, Accountancy and Finance	38	10.3	3.7	14.0
Dance, Drama and Music	5	0.7	1.1	1.8
Economics	1	0.4	0.0	0.4
Education	32	8.1	3.7	11.8
Engineering	14	1.8	3.3	5.1
English	14	2.6	2.6	5.1
Geography, Earth and Environmental Sciences	22	2.9	5.1	8.1
Health Sciences and Practice	27	6.3	3.7	9.9
History, Classics and Archaeology	10	2.2	1.5	3.7
Hospitality, Leisure, Sport and Tourism	3	1.1	0.0	1.1
Information and Computer Sciences	22	6.3	1.8	8.1
Languages, Linguistics and Area Studies	5	1.5	0.4	1.8
Law	3	0.7	0.4	1.1
Materials	0	0.0	0.0	0.0
Maths, Statistics and Operational Research	11	2.9	1.1	4.0
Philosophical and Religious Studies	4	0.0	1.5	1.5
Physical Sciences	2	0.0	0.7	0.7
Psychology	14	2.6	2.6	5.1
Social Policy and Social Work	9	1.5	1.8	3.3
Sociology, Anthropology and Politics	8	2.2	0.7	2.9
Total <i>N.B. Where a publication falls into more than one category it has been counted under each of the categories concerned</i>	272	61.4	38.6	100.0

EndNote field: context

- 3.12 The context in which the innovation took place is outlined, including for example, higher education institution, module and data analysed.

EndNote field: level/year of study

- 3.13 The study focused on literature on innovative assessment in undergraduate and taught postgraduate degree programmes. In addition to the level of study, the year of study is also shown. About two-thirds of the publications refer to a level of study, and these are predominantly concerned with courses at undergraduate level.

EndNote field: mode of assessment

- 3.14 The modes of assessment are listed as diagnostic, formative and summative (Table 3.6). None of the literature refers to ipsative assessment. Most reviewed publications (89%) allude to a mode of assessment; of these, 26% are concerned with both formative and summative assessment. Overall, summative assessment is more often discussed than formative assessment. Diagnostic assessment is only mentioned in four publications.

Table 3.6: References by mode of assessment

<i>Mode of assessment</i>	<i>N</i>	<i>%</i>
Formative	150	38
Summative	236	61
Diagnostic	4	1
Total <i>N.B. Where a publication falls into more than one category it has been counted under each of the categories concerned</i>	390	100

- 3.15 The database can be searched across references and by all the EndNote fields described in this section.

4 Analysis by themes

4.1 Introduction

- 4.1.1 This section of the report presents a set of more searching critiques of publications retrieved on six themes: presentations and other non-written assessments; portfolios and other unconventional writing assignments; assessment of groups and collaboration; student involvement in assessment; use of new technology in assessment; and what guidance and feedback is given, by what means, and when. These themes were chosen so as to be representative of the range of themes likely to be found in the innovative assessment literature. Each theme was reviewed by one member of the review team, with a second member playing a

support role by providing a check on the accuracy of his or her categorisations. It should be noted that since the final phases of work on both the database and the critiques had by necessity to be done concurrently, the numbers of publications discussed under each theme below will not precisely match those given above in section 3. Each critique follows a similar format. Concluding observations on the thematic critiques are made in the closing section of the report (see section 6.6-6.13).

- 4.1.2 The following sections were designed to be read in conjunction with the extensive EndNote database. For reasons of space, and because the database can be searched by the fields described in section 3 above, the analyses which follow focus on providing an overview of the findings rather than providing details of the relevant examples, which can all be found in the database itself.

4.2 Presentations and other non-written assessments

Defining and delimiting the theme

- 4.2.1 This theme concerns innovations in the use of presentations and other non-written assessments. Such assessments are widely believed to offer opportunities to engage students with learning, to develop communication skills and to simulate authentic situations not available with traditional written coursework and examinations. The types of assessments reviewed are listed in Table 4.1. The categorisation of an assessment type as *innovative* depended on the context. For

Table 4.1: Classification and frequency of types of non-written assessment reviewed

<i>Assessment method</i>	<i>Description</i>	<i>%^a</i>
Viva	Oral examination for a single student	22 (13)
Group presentation	Oral presentation given by two or more students	33 (50)
Individual presentation	Oral presentation given by a single student	44 (26)
Discussion	More or less structured discussion between a group of students, sometimes involving a tutor	11 (7)
Debate	Structured exchange with arguments presented for and against a position; sometimes also including a <i>jury</i> panel	0 (10)
Artefact	Production of a physical artefact, such as a display stand or non-written poster, often involving subsequent oral explanation	6 (7)
Audio/video	Production of a video or audio recording for assessment	6 (3)
Role-play	<i>Acting-out</i> of particular professions or people	11 (10)

^a Percentages are for literature classified within this theme, and (in parentheses) for all relevant literature within the database. N= 19 and 31 respectively. Percentages add to more than 100 because some items referred to more than one type of assessment.

example, the traditional PhD viva is not an innovation and is therefore not considered (although innovations in preparing for it are), while the use of vivas in undergraduate modules was included. A large literature exists on poster presentations. Most of this was not considered as it addresses written assessment. However, where posters were clearly identified as primarily or solely visual/pictorial in content, and in cases where a major part of their assessment involved oral discussion and explanation, they were included. The main overlaps of this theme with others were with *student self and peer assessment* and *group and collaborative assessments*.

Items retrieved by type or nature of innovation

- 4.2.2 Twenty-one items were provisionally placed in this theme; two were subsequently rejected as not entailing sufficient innovation. A further 12 articles were of direct relevance (in that their main topic was a non-written assessment), but were eventually classified under *student self and peer assessment* (6) and *group and collaborative assessment* (6). Most items within the theme concerned student presentations (Table 4.1). A majority of these were by individual students, but this was biased by the inclusion of many group presentations in the group and collaborative assessment theme; including these items results in 50% coverage of group presentations. While debates were discussed in the literature in other categories, none of the items classified within this theme discussed them. Most items described the assessments from the tutor's perspective (although often including some student evaluation). Four, however, were concerned principally with the collection and analysis of student views about non-written assessments. One article considered the views of potential employers (as well as academics) on communication skills. Two articles concerned fieldwork and the opportunities this presents for innovative assessments. All items addressed undergraduate assessment, with the exception of a single article considering the use of mock vivas for PhD students.

Items retrieved by subject area

- 4.2.3 A wide spread of subject areas was covered. Two articles were from each of the four areas: English; Philosophical and Religious Studies; Geography, Earth and Environmental Sciences; and Business, Management, Accounting and Finance. Single articles were from the five subject areas: Mathematics; Information and Computer Science; History, Classics and Archaeology; Social Policy and Social Work; and Engineering. Five articles were not linked to any particular discipline. Hence while many subject areas were covered, there were more articles from the arts and humanities than from traditional science subjects.

Items retrieved by genre

- 4.2.4 The largest category, with nine items, was *accounts of practice*. While items within this category all shared the defining characteristics of being practitioner descriptions of studies or innovations, most of which were relatively small-scale, they showed wide variation in content. For example, one concerned a collection of accounts from a larger project ('Assessment and the Expanded Text'), which was a collaboration between the Higher Education Academy English Subject Centre and other groups. The rest concerned generally single modules or courses at single institutions. Only one account of practice gave quantitative evaluation data; two provided qualitative data while the rest did not attempt any formal evaluation.

Empirical studies, with six items, was the next largest category. Five of these studies collected quantitative and/or qualitative data on student perceptions of non-written assessments. One study surveyed academics and employers on their perceptions of key communication skills. Sample sizes were generally small (fewer than 70 in all cases). There was a single evaluation, involving questionnaire responses from former history students on their experience of seminar presentations. This had the largest sample size (172) of papers reviewed in this category. A publication by the Staff and Educational Development Association was the single *compendium of evolving practice* considered.

Drivers to innovation

- 4.2.5 Unsurprisingly, all studies cited the need for teaching communication skills as one of the key motivations and justifications for using oral and non-written assessment. Most also discussed other key or transferable skills that the authors thought were encouraged by non-written assessments. These were usually related to particular types of assessment; for example, group presentations were often seen as facilitating group working skills. The need for these skills was identified from external agencies (e.g. QAA), from employers and from the experience of staff. Most studies also referred to the literature on deep versus surface learning and implied or stated that oral assessment methods are likely to foster the former and lead to better engagement of students with assessment tasks. A common theme was the notion of a public performance being particularly motivating (as well as scary) for students, and the idea of openness and interactivity in questioning leading to critical understanding. In some cases authors suggested that endorsing the skills agenda (e.g. teaching communication skills), will allow them to retain support for teaching and learning methods, such as fieldtrips, which are becoming harder to sustain given large class sizes and falling resources. A number of the studies discussed student involvement in the negotiation and setting of assessment criteria, as well as the marking of the presentations. Hence a desire to involve students in assessment, sometimes explicitly stated as *democratising assessment* or as changing power relationships, was also a motivation.

Points of reference

- 4.2.6 Literature cited as informing or inspiring the studies considered, varied widely. However, four broad categories, each informing two or more studies and none mutually exclusive, were identifiable. Six of the studies referred to the literature on engaging students in their learning, encouraging active rather than passive or deep rather than surface learning; literature cited included the work of Ramsden, Entwistle and Race. A second point of reference, particularly important in two of the studies, was the rather small body of conceptual work on oral presentations in higher education, with Joughin a key influence. Many of the studies cited subject-specific literature, including research papers and publications by professional bodies, such as the Council for College and University English and the Higher Education Funding Council for England and Wales, recommending the fostering of particular skills within subject areas. The fourth category providing a point of reference was reports of previous experience by practitioners of using the various assessment methods and of students experiencing them.

Implications drawn by authors

- 4.2.7 The importance of clear criteria was raised in many of the papers; partly as simply good practice, but also as a reflection of the wariness over the possible subjectivity and unreliability of oral versus written assessment. Some work discussed student fear and apprehension before oral assessments, and implied that these are scarier than other types of assessment. This would suggest the need for low-risk training and practice in the necessary skills, although it is not clear whether oral assessments are really scarier or just more novel (the few studies considering student perception of oral assessments supported the idea that they are scarier, but also suggest that anxiety diminishes with practice). A number of studies identified the *public* aspect of oral assessments, such as presentations before peers, as crucial in fostering greater engagement with the task (but also in adding to the stress). Hence there is an implied need for the right physical space for peer observation and assessment. Some of the types of assessment, particularly one-to-one vivas and assessed discussions, are likely to be very time-consuming and thus restricted to relatively small classes. One paper discussed the implications of using oral assessments for students with special educational needs and from minority backgrounds, with the suggestion (supported by some evidence) that such students will perform better in these assessments than in more traditional written ones.

Concluding observations

- 4.2.8 Accounts of practice, generally written by subject-specialist academics, was the dominant category, and the quality of studies considered varied widely. Some provided excellent overviews of innovations in particular contexts, which stimulated reflection, gave good ideas transferable to other areas and gave a feel for the challenges involved in implementation (see for example *Gent, Johnston and Prosser, 1999).¹ However, others were frustratingly vague, omitting basic details about sample sizes and populations. In general these studies did not attempt any formal evaluations of the efficacy of the innovations described; only one contained a quantitative evaluation, with two more giving qualitative data. Hence while most studies claimed important benefits for the innovations described (perhaps reflecting the tendency of such work to be written by practitioner enthusiasts), it was generally difficult to place such claims in the context of the costs of the methods and their strengths and weaknesses compared with other assessments. This might reflect the ambiguous and flexible nature of accounts of practice, which lack clear traditions of content and standard. However, similar conclusions apply to the empirical studies considered, largely because they had small sample sizes and were usually limited to collecting student perceptions only. One study (*Jackson and Henry, 2003) provides an interesting comparison between performance in oral and written assessments, although again this was based on a small sample. Thus there is a need for larger studies, preferably involving proper controlled comparisons of different methods (e.g. oral versus written assessments) or longitudinal work studying the development of students' skills and perceptions.
- 4.2.9 While it is trivially true that oral assessment methods are likely to improve oral communication skills more than written assessments, it does not follow that they

¹ * refers to a piece of literature that is included in the EndNote database but not listed in the references at the end of this report

will also engender deep learning, as was often claimed in the literature reviewed. It is likely that methods that allow more open dialogue (discussion and vivas) will do this better than more closed unidirectional methods (presentations). Again, this would be a useful area to research. While the emphasis in the literature was often on the development of a range of transferable skills, it is also possible that oral assessments are better at reaching and testing *traditional* academic outcomes (such as theoretical understanding) than written ones. This would follow from the notion of greater student engagement with oral than written assessment, and deserves greater attention. Some literature (for example *Jackson and Henry, 2003) suggests that oral assessment is likely to be fairer for students with special educational needs (e.g. dyslexia) and for ethnic minority students and students from non-traditional backgrounds. This interesting suggestion, along with the general question of the benefits and disadvantages of non-written assessments to students with different personalities, aptitudes and learning styles, deserves further research.

4.3 Portfolios and other unconventional writing assignments

Defining and delimiting the theme

- 4.3.1 This theme is concerned with portfolios and other unconventional writing assignments such as learning journals, project documentation, diaries, reading logs and reflexive notebooks. Such written assessments are used to record reflections, build evidence for learning and demonstrate development of learning. In all of the papers that were reviewed, the opportunity to reflect upon experience to promote deeper quality of learning featured strongly. The main areas of overlap of this theme are with *student involvement in assessment* particularly self-assessment and to a lesser extent with *guidance and feedback to students*. It also links with another sub-theme not dealt with in this review: *students on placements/in the workplace/in situ*. E-portfolios did not feature in this collection of material, despite the evidence from Medicine of increased use (Cotterill, Bradley and Hammond, 2006).

Items retrieved by type or nature of innovation

- 4.3.2 A total of 40 items were retrieved and judged to be relevant to this theme of unconventional writing assignments (i.e. those that departed from the traditional essay type of assessment). Of that number of papers gathered, ten specifically referred to portfolios, five referred to learning journals, a further six were characterised as reflective or reflexive journals and the remainder were explained as projects, diaries, log books and different forms of textual practices. While nursing and teaching featured strongly in the use of portfolios, other disciplines were using this form of assessment as a means to improve writing skills, the development of argument or to triangulate assessment methods. One of the papers, *McMullen et al (2003), offered a literature review of portfolios and the assessment of competence in nursing in the period 1989-2001.

Items retrieved by subject area

- 4.3.3 The two largest subject groupings that featured in this theme were Education (9) and Health Sciences and Practice (9) but when other health or social care aspects

were added that accounted for Sociology, Anthropology and Politics (2); Bioscience (2); and Social Policy and Social Work (2). English accounted for five; Geography, Earth and Environmental Sciences (5); and the rest were Engineering (1); Information and Computer Science (1); Mathematics, Statistics and Operational Research (1); with the remainder from non-specified areas. Given the attention that work-based assessment receives within vocational education, it was surprising that there were not more papers from Health and Social Care.

Items retrieved by genre

- 4.3.4 With respect to identified genre the greatest number (28) were judged to be accounts of practice. There were four enhancement projects, three theory/conceptualisation papers, two empirical studies, one evaluation, one set of guidelines and twenty reviews of the literature. A number of overlaps of genre occurred particularly between an account of practice and evaluation or empirical work. Of the two empirical studies retrieved, both used qualitative methodology one in the form of an illuminative review in nursing and the other study sought to explore two Open University courses – one of which used closed book examinations and the other used a written project as an end-of-course assessment. The one evaluation project was concerned with role of foreign fieldwork in promoting deep learning.

Drivers to innovation

- 4.3.5 The impetus to use portfolios, learning journals or diaries was related to principles of lifelong learning, active learning and in particular this form of assessment provided a place reflection on experience. This form of assessment was process-driven and considered to foster metacognitive skills, self-directed learning and to integrate learning. As such, it was believed to endorse adult learning principles. Portfolios did feature as one of the LTSN Generic Centre's series of assessment guides and briefings and in at least three papers the drivers for using portfolios was directly prompted by their professional body to determine evidence of self-directed learning.
- 4.3.6 There is also a detectable shift in the literature from an accent upon *reflection-on-practice* to a more critical intellectual ability to be reflexive in accounts of experiential learning. This was linked in one case by the wish to counteract the perceived reduction in dialogue between lecturers and students where the portfolio was used to provide evidence of depth and breadth of understanding. Overall, portfolios, learning journals and diaries provide a source of written evidence, which encourage critical reflection and in some instances offer the opportunity for self-assessment. Integration of assessment within learning is a key feature of project work from which learning logs are an outcome. The possibility that improving writing and documenting skills were transferable skills and necessary for employability was also mooted. Assessment of portfolios can be problematic as Falchikov (2005) has noted and a number of papers endorsed the idea that clear criteria referenced frameworks were necessary to underpin the process. Only two papers begin to question the marking of portfolios and while *Baume and Yorke (2002) questioned the inter-rater reliability of markers, *Webb *et al* explore how portfolio assessment processes should be evaluated. Earlier, *Morrison (1996) sought to problematise the whole area of reflective practice.

Points of reference

- 4.3.7 Conceptually, constructivist ideas about learning and principles of lifelong learning underpin the use of this approach to documenting learning, along with references to student-centred learning and autonomous learning. Equally, ideas that reflection helps to integrate theory and practice are also prevalent regardless of the subject area that is being explained. Building assessment into the process of learning, and providing a place for students to take ownership for their learning alongside growing awareness of the process, is a key point of reference within the papers in this section. Empirically, the intention of most of the papers was to share good practice or to illuminate issues that emerge from marking such lengthy pieces of work. Some sharing of knowledge across the papers was revealed in reference lists but this did not include Birenbaum's (1996) rubric for judging a portfolio. *Platzer, Snalling and Blake (1997) in the context of nursing pointed out the paucity of studies in evaluating the place of learning diaries or journals within assessment practice. Key writers who are quoted in most papers are – Kolb, Boud, Keough and Walker, Schon, Mezirow, Crème and Baume and there are profession-specific writers, such as Johns, who are always mentioned within the nursing literature.

Implications drawn by authors

- 4.3.8 There are logistical issues concerning the use of portfolios or longer pieces of writing in assessment which are concerned with agreement on how to structure portfolios, confidentiality issues and how to manage the more personalised aspects of learning, how to give feedback promptly, and how to achieve inter-rater reliability. Current resource issues within higher education make it challenging to deal with such large quantities of written material despite the benefits which are believed to accrue. The consensus about the advantages range from: portfolios or learning journals providing documentation of learning over time; providing a record of learning and a resource for future learning; improving written communication skills; and subscribing to the values of lifelong learning. In the papers which were reviewed, a very small percentage were using assessment of this kind only for formative purposes, while 17 of the papers noted that this kind of assessment was used for formative and summative purposes. Another 17 noted that this kind of work was used for summative assessment and this ranged from 10% - 50% of the overall weighting. Staff training would seem to be necessary to improve practice and to ensure that issues of rigour are addressed. Equally, the opportunities for self-assessment or peer-assessment are raised but not extensively debated within the papers selected for this section. E-portfolios (as outlined in an exploratory way by Cotterill, Bradley and Hammond (2006) within the context of Medicine, and which allow for the possibility of accessing portfolios on the internet and providing both formative and summative feedback) did not feature within the 41 papers reviewed.

Concluding observations

- 4.3.9 While portfolios and learning journals are commonly a feature of vocational education and in particular workplace learning in either nursing or teaching, in this corpus of literature there are many examples of innovation within discrete areas. For example, *Townend (2001) in the area of engineering mathematics and *Ovens (2003) in the area of an imaginative use of writing called *patchwork text*. Most work was extensively referenced and offered a case example or an outline of

practice that could be replicated. The papers often represented the first attempt to make public an innovation in learning, teaching and assessment and as such predated more empirical studies. There is considerable opportunity for more work on issues such as commonality of structure, agreed criteria for content, what type of formative assessment is useful and study of the inter-rater reliability of marking of summative work. Alignment between the chosen philosophy of learning, teaching approaches and assessment approaches is evident in the papers. Within health and social care, regulatory bodies now require practitioners to demonstrate continued professional learning through the documenting of evidence within a portfolio. This forms part of transparent approaches to professional development and most universities have sought to embed this within undergraduate and postgraduate curricula to ensure conceptual continuity. Interestingly, while professional journals may contain more material related to this form of assessment, recourse to publishing within education journals is less and in this body of work only accounts for only five papers.

4.4 **Assessment of groups and collaboration**

Defining and delimiting the theme

- 4.4.1 This theme concerns the assessment of students working collaboratively in groups of varying sizes and in different kinds of ways. Students may work on projects leading to reports, posters and/or presentations. They may work on-line or face-to-face in groups, which are structured and managed to greater or lesser degrees. In some contexts the students are given a project or problem and can allocate tasks between them as they wish, while in others they may each be required to help with the completion of every part of the project (for example, the report, the presentation etc.). The projects may involve fieldwork or work-based learning. The students may have to search for literature, interview staff or solve set problems. In some cases the students work on individual assignments and the collaboration consists of having drafts reviewed by other students. This theme overlaps with several of the other themes discussed in this review. The strongest overlap is with *student involvement in assessment*: the use of peer-assessment is considered in 11 of the 27 articles reviewed for this section. The articles also overlap with the *use of new technology, guidance and feedback, presentations and other non-written assessments, portfolios and other unconventional writing assignments and assessment of students in the workplace*.

Items retrieved by type or nature of innovation

- 4.4.2 While some of the innovations reported involve the setting up of collaborative or group assessment, and some provide advice to those planning to do this, the innovations tend to be in altering established practices or finding out more about their effectiveness. One article extends and adapts the practice of assessing group work to include first-year students (e.g. *Bourner, Hughes and Bourner, 2001), while another is concerned with using group assessment to help students develop skills which they need in their transition from school to university (*Booth, 2001). One innovation was introduced in order to help first-year students understand the nature of research being carried out in their department, while getting to know both staff and students (*Dwyer, 2001). Some of the innovations concern the introduction of work-related group assignments and/or the

involvement of employers in assessing these projects, in order to develop transferable skills (e.g. *Goldfinch, Laybourn, MacLeod and Stewart, 1999). In others group assessment is seen as helpful for developing learning skills or to instil in students the notion of a learning culture which is collaborative rather than competitive. Four of the articles focus on assigning students to groups, whether randomly, by student choice (friendship groups), or by deliberately selecting students to create mixed-ability, streamed or multi-cultural groups – and the effect of this on the grades the students receive. Several of the articles concentrate on the issue of assigning marks either to individuals or groups or to a group moderated by adding or subtracting marks for individual contribution (often by peer-assessment). Two of the innovations are specifically on training programmes for students to help them develop group skills, and this focus on training is part of other innovations reported. In two of the articles the innovation reported is to make the collaboration on-line (*McConnell, 2002, *MacDonald, 2003), thereby providing a transcript of the way students are working together. Few of the innovations extend beyond one course or department, although a few do survey other departments in their subject area to compare practice (e.g. *Lejk, Wyvill and Farrow, 1997).

Items retrieved by subject area

- 4.4.3 There are more articles about innovations in the assessment of groups and collaboration in the social sciences than other areas (Business, Management, Accountancy and Finance (5); Geography, Earth and Environmental Sciences (5); Education (2); Sociology, Anthropology and Politics (1); Hospitality, Leisure, Sport and Tourism (1); and Social Policy and Social Work (1)). Five of the articles were linked to Information and Computer Sciences; with one each in Mathematics, Statistics and Operational Research; Bioscience; Built Environment; and Engineering. There were few innovations reported in the arts and humanities (English (1); History, Classics and Archaeology (1); Dance, Drama and Music (1); and Philosophical and Religious Studies (1)). Most of the articles concentrated on only one subject area and three were not specific to any one subject area. However, it should be noted that, while most of the articles were in education teaching and learning journals, the five Geography articles were from the *Journal of Geography in Higher Education*. One of the Business, Management, Accountancy and Finance articles, one of the Information and Computer Sciences articles and the Philosophical and Religious Studies article were also from subject-specific journals. There was one book chapter.

Items retrieved by genre

- 4.4.4 By far the most common genre used in this theme was that of an account of practice, with 16 of the 27 articles reviewed being categorised as this. There were also six empirical studies, one paper that was classed as theory/conceptualization, one report of an enhancement project, one evaluation, one article that offered guidelines and one review of the literature. However, there was some overlap between these genres, with many of the accounts of practice presenting or referring to some data. The guidelines article refers to case studies and the review of the literature also includes an element of an account of practice, with practice in two of the authors' department being described along with practices referred to in the literature. Where data had been collected, it was usually through a questionnaire asking students for their perceptions of the innovation in assessment. However, in five cases students had been interviewed, taken part in

focus groups or had sent email responses. One article used data from course evaluation questionnaires over several years. In three articles staff had completed questionnaires and/or been interviewed, and in one of these articles employers who had taken part in the assessment of groups were included. On-line transcripts of student collaboration were analysed and reported in one instance and four articles analysed student grades. Two articles surveyed practices across several departments (in one case across the UK for one subject area). Ten authors reported collecting data by more than one means. In six of the articles, no data was collected but authors' perceptions or those of other staff collected in an ad hoc manner were reported.

Drivers to innovation

- 4.4.5 The driver for the innovation is often a perceived need to produce graduates who have transferable skills, including the ability to work in teams and collaborate, as these are skills are considered to be valued by prospective employers. In some cases the students are taking vocational or professional courses, where group work is seen as not just desirable but necessary for their career, and it is seen as important that what or how the students practise at university matches working practice. In other cases the driver may be the departmental or institutional context: the need to cope with an increasing number of students or students from more diverse backgrounds (including from overseas) and to enable students to benefit from the experiences of those of other cultures. Group work may also be used as a way of helping students make the transition to higher education, by encouraging socialisation and the discussion of what it means to study in a particular discipline. In several of the articles the driver is an aspiration to encourage students to see the learning environment as a collaborative one rather than a competitive one. In some instances the consideration is to develop students' learning or to find out whether students learn best through group or individual assessments. Collaboration and group assessments may help students to see new perspectives and provide a fuller consideration of issues, helping them to develop their ability to understand, analyse and interpret. Self- and peer-assessment is often seen within this context as helping students to reflect on their experiences of group work. And providing an assessment element may encourage participation in group tasks. Often the group assessment itself is not an innovation but an aspect of it is, or the authors want to explore whether, for example, grades given for group work favour weaker students, or whether selecting groups in different ways affect group marks and/or enhance the achievements of students. In one case the driver for carrying out a survey was to find out how widespread the practice of group assessment was within a discipline and what academics saw as the issues to be considered.

Points of reference

- 4.4.6 There was no one body of literature that appeared to be generally referred to by authors in writing about group assessment. The points of reference varied a great deal. Reflecting the drivers for innovation, reference to the transferable skills agenda was common and used to set the context for an innovation. Perhaps because the authors were in many cases academics within disciplines other than academic development, the references to the student learning literature were, while reasonably common, usually very selective. They included literature on approaches to learning, the pedagogy of on-line learning and teaching and learning styles. Handbooks on how to teach in higher education were often the

source. Theories and concepts such as Kolb's learning cycle, Wildermeersch and Jansen's double-loop learning concept or Honey and Mumford's learning styles were also used (in one instance each), but were referred to as isolated concepts with little reference other possible theories. Perhaps surprisingly for practising academics, the references were not only from a narrow spectrum, but were sometimes hardly in evidence. In a few cases the article had been published with only three or four references of any kind. Commonly the literature was only referred to in the introduction to an article, with no attempt to relate the article's findings to the literature. Previous research on the specific aspect of assessment being discussed in the article, by the same or different authors, was used as a point of reference for a study in a few cases. Only about a third of the articles referred to literature on assessment, whether group, peer- or self-assessment or more generally. In a few instances the main point of reference was an issue or issues raised by students or staff, with little reference to any previous research in the area.

Implications drawn by authors

- 4.4.7 A wide range of issues are raised by the articles reviewed. In most cases, the authors of the articles are positive about their own and their students' perceptions of the usefulness of assessing groups and collaboration in developing a range of skills, including the ability to cooperate as a team, to think critically and to reflect on their learning. Giving grades for group work was considered important if students were to take the task seriously. Authors identified many practical issues for academic staff to address, including: course organisation and management; preparation and training of both staff and students; the amount of support needed by staff to facilitate the group learning; the need to match tasks to learning outcomes; organising tasks so that each student develops a range of skills; and dealing with problems within groups (such as passengers) or problems with finding times for groups to meet. The usefulness of using e-learning was only raised in the two articles where the course described was studied on-line.
- 4.4.8 One widely-raised issue was that of how to allocate marks to groups and in what sense marks could be considered 'fair' for any individual student. This had implications for how students were allocated to groups, whether by random or selective means, with opinion divided on the effect and importance of this. There was also disagreement between the articles on how far group work might benefit weaker students or hinder brighter students, and whether that mattered. The issue of what was being assessed, whether it was the product or the process or both, was considered and whether the skills needed to produce good group assignments were as valid as those that produce good individual work. The importance of making both the criteria for marking and the purpose of the assignment clear to students in order for them to see the value of group assessment was stressed in several articles.
- 4.4.9 Some of the issues raised went beyond a single course. Some of the articles urged a change in culture in higher education, with students being encouraged to see learning as a collaborative rather than a competitive enterprise. Departmental and institutional policies would be needed in order to implement some of the suggestions, such as coordinating assessment so that students have enough but not too much group assessment, or so that group skills are developed during a student's time at university.

Concluding observations

- 4.4.10 The articles reviewed varied a great deal in several respects. Some articles set out to provide advice for practitioners based on the author's personal experiences and perceptions. Others set out to research a particular issue, using data either from the author's own students or collected more widely. Many articles simply reported an innovation being carried out in their department, and on a course on which the author taught. Some of these articles provided a clear rationale, a basis in the literature, and a clear presentation of results and their implications (e.g. *Knight, 2004, *Huxham and Land, 2000). Others might benefit from a greater awareness of research that has been done previously or other relevant literature, and in some cases a more systematic analysis and presentation of the data would make the robustness of the research clearer to the reader. The sample sizes are small in almost every case – data from just one course unit, often collected only during one academic year – making it difficult to generalise from the results. However, the articles do raise a large number of implications and issues, based on a wide range of experiences, which will be useful for academic staff involved in assessing groups and collaboration.

4.5 Student involvement in assessment

Disciplinary areas represented in the student involvement literature

- 4.5.1 A wide variety of disciplines are using and researching ways of involving students in assessment. The largest concentration of entries is found in the area of Biosciences (17%), though it should be noted that five of these were by a single research group (*Orsmond et al.), whose publications documenting the development of their investigations into the type and use of criteria in self- and peer-assessment span the period between 1996 and 2004. Next greatest concentrations occur in four areas – Business, Management, Accountancy and Finance; and Information and Computer Sciences (14% each); and Education; and Geography, Earth and Environmental Sciences (10% each).

Overlap with other categories

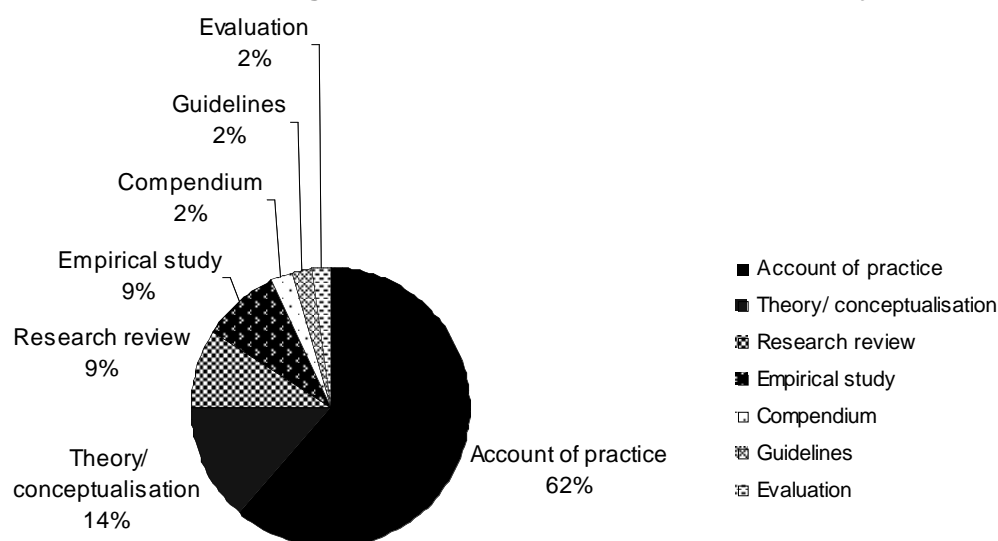
- 4.5.2 Not only is the *involving students* category literature found across the disciplines, it may also be seen as occupying a central position in relation to other thematic categories. It has the potential to overlap with many other categories: the use of technology; portfolios; group work; oral presentations; and feedback. In turn, technology has been found to overlap with portfolios; portfolios with groups; groups with oral presentations; presentations with feedback; feedback with technology and groups. The literature on involving students in assessment illustrates many of these overlaps. For example, 31% of entries that reported work on involving students in assessment featured groups and group work. A further 18% dealt with aspects of feedback, 13% with presentations and 9% with posters. Computer Aided Assessment (CAA) featured in 3% of studies involving students, and there were single instances of overlap with preparation of diaries and performance. 20% of entries reporting involvement of students in assessment exhibited no overlap with other categories.

Genre of study

4.5.3 As indicated in Figure 4.1, nearly two-thirds of all papers categorised as reporting the involvement of students in innovative assessment are accounts of practice. The genre of study is related to the intended audience for the paper, with accounts of practice, guidelines and compendia aimed predominantly at practitioners. Papers with a greater theoretical or conceptual emphasis are of potential relevance to researchers or policy makers, though practitioners may also benefit from reading them. Similarly, reviews of the literature are of equal interest to both practitioners and researchers. Empirical studies also seem to be aimed at either audience. The type and quality of entries in the *accounts of practice* category varied widely. At best, entries reported well-designed studies thoroughly, rigorously and clearly (e.g. *Fitzpatrick, 2006; *Jordan, 1999; *Pain and Mowl, 1996). At worst, accounts of practice reported ill-conceived and poorly-designed studies badly, such that replication was impossible, and in which conclusions drawn went well beyond the data, supporting evidence being absent. 14% of the literature was categorised as *theoretical or conceptual*. Again, quality of entries varied, though the standard was more even than in the accounts of practice category. The theoretical entry by *McLuckie and Topping (2004) was deemed particularly fit for purpose, as was the book by *Falchikov (2005).

Figure 4.1

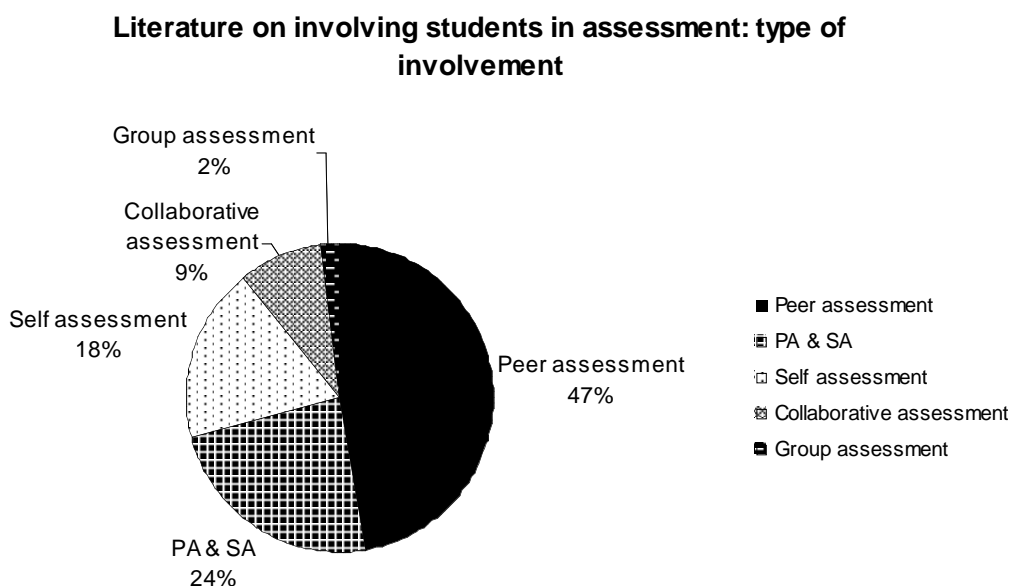
Literature on involving students in assessment: Genre of study



How students are involved

4.5.4 As indicated in Figure 4.2, peer-assessment was the most frequently occurring type of student involvement in assessment within the literature, which may not be too surprising, given the overlap between studies of student involvement and studies involving group activities. Nearly half the entries reviewed (47%) involved students assessing their peers, and a further 24% involved both peer- and self-assessment. Self-assessment featured in 18% of entries in this category. While

Figure 4.2



collaborative assessment and group assessment were also found, they constituted a relatively small overall proportion. Of course, there may be little consistency in defining *collaboration* or *co-operation* or in labelling of studies by authors themselves and mistakes in categorisation within the present survey. However, it can be argued that much peer-assessment, particularly when it involves groups, is also collaborative. There was some variation within the group of peer-assessment studies, regarding focus and rationale. Half the peer-assessment entries related to peer-assessment of a group. Two-thirds of these applied to assessment of the group of which a student was member and the remaining third to assessment of another group. A further 19% of this sub-set of entries applied to peer-assessment of individual products or performances. The remaining 31% of studies involved peer-assessment which was not elaborated further.

Drivers to innovation

4.5.5 Authors are stimulated to carry out and report their work for a variety of reasons: potential benefits to students; in order to investigate and solve problems identified with an innovative practice; as a response to external or internal drivers; or to benefit themselves, their teaching and that of their colleagues. As indicated in Figure 4.3, *benefits to students* constitutes the main reason authors designed, conducted and reported their work, accounting for a third of all entries. As well as general unspecified benefits, individual benefits reported include:

- empowering learners
- encouraging attention (relating to oral presentations), dialogue/creating partnerships, more proactive strategies in teachers and students, reflection
- enhancement of learning, performance, personal development/autonomy

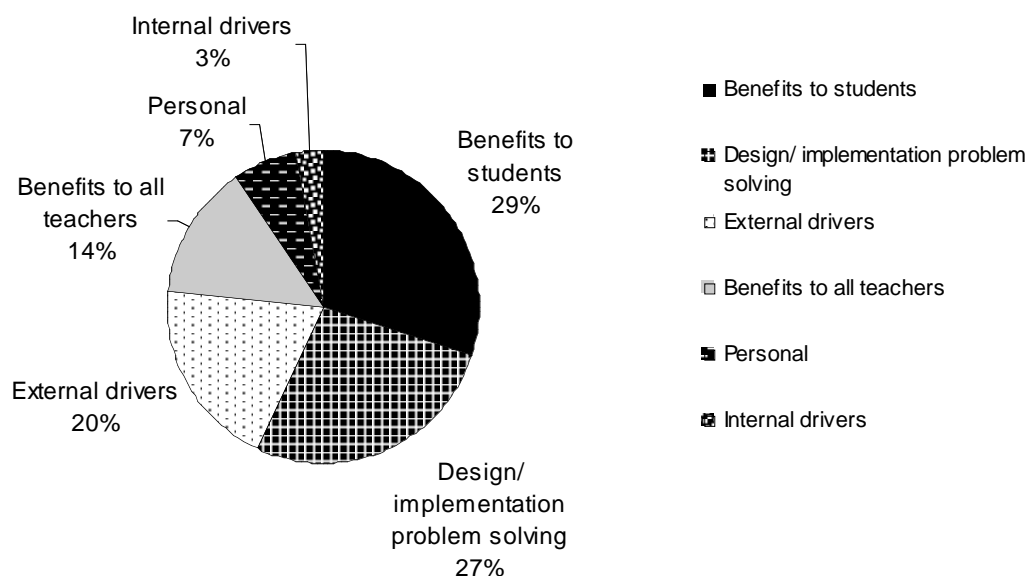
- improvement of affect, student understanding of assessment, student understanding of criteria.

4.5.6 Over a quarter of authors report their attempts to solve problems relating to the implementation of the innovative assessment practice itself, identified by either themselves or others. For example, *Taras (2002), in her study of feedback, aimed to investigate the contradictory messages we may be sending to our students. *Gale et al. (2002: 558), on the other hand, set out with a "genuine desire to examine the nature of assessment", while *Penny and Grover's (1996) paper arose out of a wish to investigate paradoxes. Also included in this category are reports of tests of reliability and validity of practices, as well as the development of algorithms. A few papers (those overlapping with assessment of group work) investigate the problem of how to differentiate between individual members of a group.

4.5.7 External drivers to innovation included not only specific named bodies, but also the recent expansion of higher education, the needs of the global economy, the needs of the learning society, or the work of others. Internal drivers included the increased use of on-line or distance learning in higher education and feedback from students. Benefits to teachers included: the encouragement of dialogue and creation of partnerships; encouragement of more proactive strategies; improvement of assessment and improvement of teacher understanding of assessment; improvement of teaching; provision of resources for teachers; stimulating thought; and increasing teacher enjoyment.

Figure 4.3

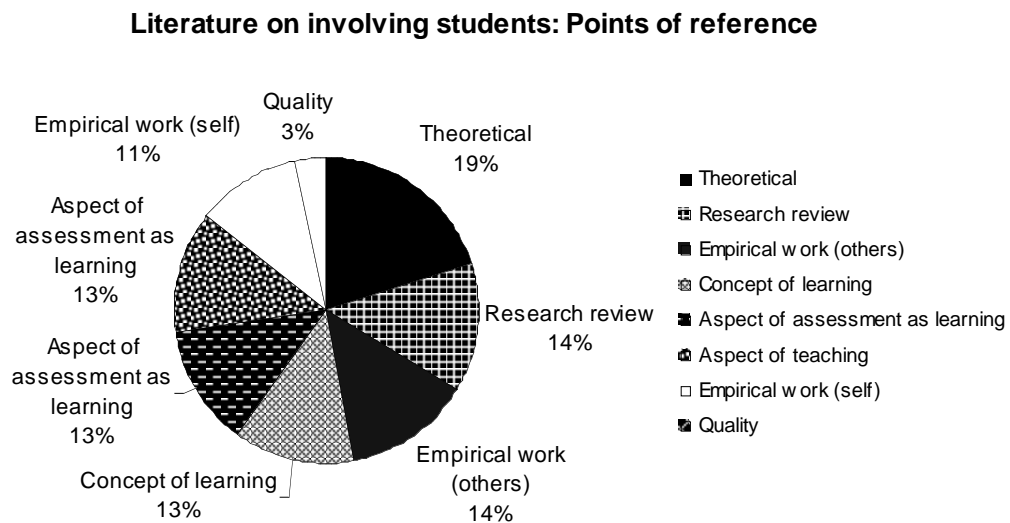
Literature on involving students: Drivers to innovation



Points of reference

4.5.8 Studies reviewed had a variety of starting points: theoretical; conceptual; those concerning aspects of teaching and learning; as well as practical points of reference. Over a quarter of studies built on the work of either the author or other researchers. A wide variety of theoretical standpoints were invoked such as constructivism, critical theory, feminist theory, Marxism, post-structuralism and communities of practice. In addition, papers were contextualised in clear areas: social; cognitive; meta-cognitive; and affect. Proportions of theoretical and other points of reference are illustrated in Figure 4.4. It is heartening to see so many authors grounding their work within a theoretical framework, although it should be noted that mention of a framework in the introduction does not always translate into a thoughtful discussion of outcomes within it. Those authors using reviews of the literature as their point of reference seem to be better placed to serve the reader, in that the reviews themselves integrate findings from many studies. In addition, links between the literature and current study are usually made. Previous work by others is a common starting point for many entries. The following authors were cited by more than one author as having influenced their innovation: Black and Wiliam; Boud; Edwards; Falchikov; Fry; Giroux; Goldfinch; Heron; Lave and Wenger; Lejk; Piaget; Rowntree; Sadler; and Topping.

Figure 4.4



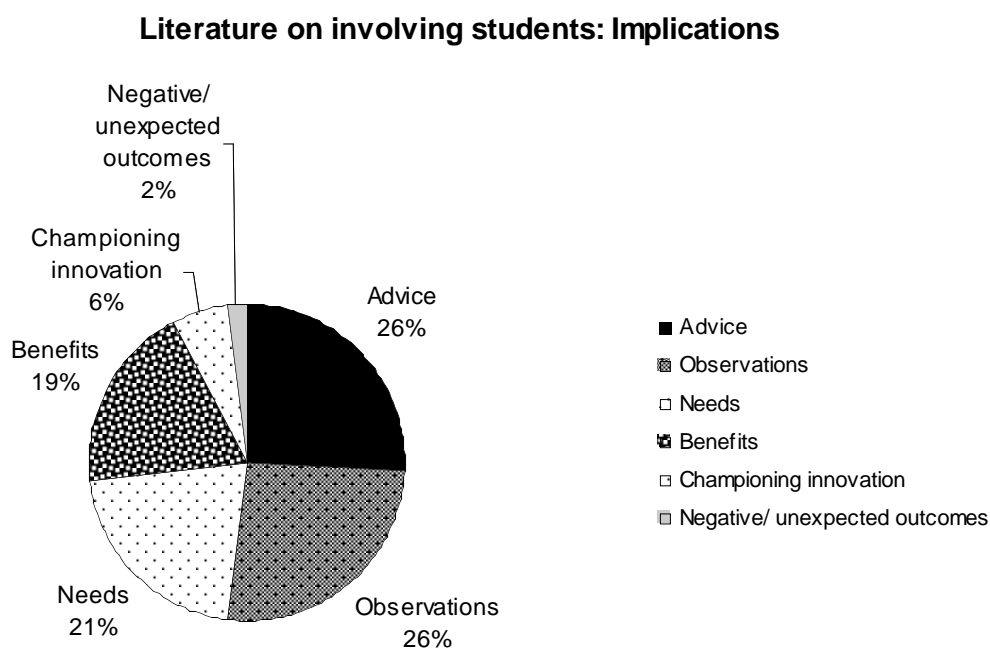
Implications drawn by authors

4.5.9 Figure 4.5 differentiates between *advice* and *observations*, the former being clearly derived from, and supported by, experimental findings; the latter from what appears to be impressions. However, as it is impossible to go beyond the author's words, this distinction may be less useful than appears at first sight. However, while some observations provide practical information for readers, they do not take the form of concrete suggestions. The most frequently occurring advice offered by authors concerned the need for preparation and training for both students and teachers. For example, *Topping *et al.* (2000) advocated the use of anonymous reports from previous cohorts as "useful desensitisation and training" for their

students before they engaged in assessment of academic writing. A similar solution was suggested by *Maguire and Edmondson (2001) for students involved in self-assessment of group projects. *Elwood and Klenowski (2002) recommended steps necessary to ensure shared understanding of assessment practice. Several authors stressed the need to make one's reasons for involving students in assessment clear, some arguing that it is important for teachers to ensure that their valuing of the system is explicit. Others recommended an incremental approach or the use of pilot studies as preparation for full implementation.

4.5.10 Some observations are also derived from the results of the study. A few authors deemed their implementation time-consuming, but some expressed the belief that the time spent was worth it. Other observations were either very general in nature (e.g. "peer-assessment promotes learning" or "computer-aided assessment looks promising") or very specific to the context (e.g. "attendance improved" or "holistic peer assessment better supports the aims of group assessments than category-based"). It would be helpful to both authors and readers for many observations to be expanded and translated into advice, for example, by grounding them in empirical data and linking them to a theoretical framework.

Figure 4.5



4.5.11 Needs identified by authors provide very useful information for both the authors themselves, and for readers more widely. About a third of what we are terming *implications* of study outcomes concerned the author identifying the need for further study, more development and evaluation. For example, *Searby and Ewers (1997) reported plans to extend their implementation of a peer-assessment scheme in one area of music into new areas. Needs also included problems and tensions in the implementation that require modification to the procedure, or interesting results that suggest further investigation (e.g. gender differences). More

practical needs are also identified: the need for staff support, and for support materials. The need to create a supportive atmosphere to encourage students to air their concerns was also identified. Time allocated to the implementation and experience in carrying it out were also seen as important.

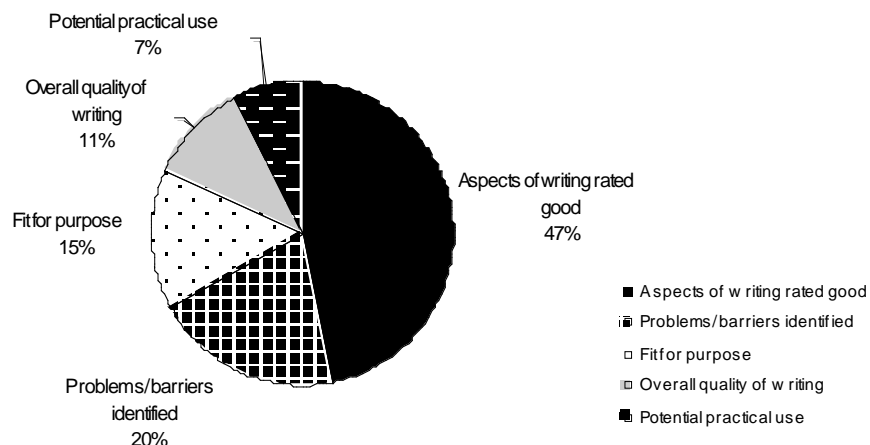
- 4.5.12 Benefits identified are similar in origin to observations, in that they tend to be very general and often unsupported by hard evidence. Some authors claim their implementation to be beneficial in aiding the development of personal and lifelong learning skills. Others claim students evaluate the experience as having benefited their learning. Other benefits claimed include helping students overcome unrealistic expectations, helping shared understanding and encouraging integration.
- 4.5.13 *Championing the innovation* is done by five authors who call to their colleagues to begin or continue assessment reform and to value it; urging them to experiment with multiple approaches, including, of course, student involvement.
- 4.5.14 Two negative or unexpected outcomes were reported. Joint construction of criteria was found to have no marked effect by one team (*Langan et al., 2005), while secret peer-assessment was found to reduce the feedback to students (*Lejk and Wyvill, 2001). Such information is useful in that it stimulates thought in the authors, researchers and readers more generally.

Critiques

- 4.5.15 An almost equal number of positive and negative features were noted relating to the reporting and design of studies. These are summarised in Figures 4.6 and 4.7. Generally speaking, over half the studies were rated as having either an overall good quality of writing or features rated as good (e.g. literature review thorough, providing a good context for the innovation). The needs discussed above may also be seen as positive features of studies. Improvement cannot be made in the absence of critical evaluation.

Figure 4.6

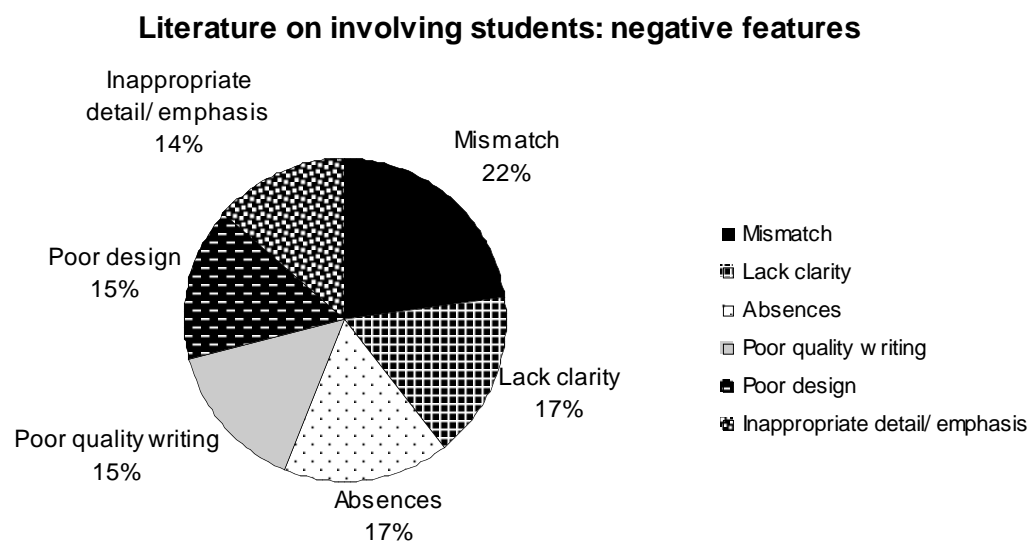
Literature on involving students in assessment: positive features



- 4.5.16 Positive features fell into roughly five categories: aspects of writing rated as good; identification of problems or barriers; being fit-for-purpose; overall quality of writing; and potential practical use.
- 4.5.17 *Aspects of writing rated as good* (47%) included features such as a sound review or synthesis of recent research which is well referenced; clear description of implementation, such that replication is possible; inclusion of exemplars; inclusion of quotes from interview or questionnaires to illustrate points; good definition of terms; results supporting conclusions; a rich, useful discussion; and exploration of issues or caution concerning claims. In addition, some entries contained an honest account of the author's own assumptions. There was one example of a paper co-written with students (*Topping et al., 2001).
- 4.5.18 *Identification of problems or barriers* (20%) covered a wide variety of topics, ranging from the acknowledgement that students might have been influenced by the role and position of the teacher and that his opinions might have been given *greater legitimacy* than those of peers by students (*Greenbank, 2003), to recognition of the conflict that exists between self assessment for summative and learning purposes (*Fitzpatrick, 2006). Some authors considered contamination, bias and threats to validity. Others documented the lessons learned and proposed changes to future implementations. One author identified the need for underpinning pedagogical principles.
- 4.5.19 Those entries rated as fit-for-purpose (15%) succeeded in supporting and advancing assessment practices and research. A good example from this category is a paper on self- and peer-assessment and feedback in the context of essay writing (*Pain and Mowl, 1996). This paper presents a clear account of the preparation for the exercise, along with examples of materials used, criteria sheet and peer-assessment form. It contains a thoughtful discussion of the results, a critical analysis and suggestions for modifications to the scheme. It must be rated as particularly helpful to those wishing to implement a similar scheme.
- 4.5.20 The rating *overall quality of writing* indicates a clear, thoughtful entry; one that is well grounded and focussed on the intended readership. 11% of entries fell into this category.
- 4.5.21 The final category of entries under this heading, *potential practical use*, was applied to 7% of positive entries. Examples include *Orsmond's (2004) case studies, from a variety of UK sources. These are presented in a consistent and clear house style. Sections on background and rationale are included, along with sections entitled 'How to do it', 'Troubleshooting', 'Does it work?' and advice on using the approach. *Falchikov's (2002) chapter on peer-assessment in '*Assessment: case studies, experience and practice from higher education* is an account of peer-assessment implementations that is likely to be useful as a staff development aid, as the chapter includes questions and invites readers to solve problems. Similarly, *Falchikov's (2005) book on improving assessment through student involvement provides a useful resource for researchers and practitioners alike. *Elwood and Klenowski's (2002: 251) account of a one-day workshop also has potential usefulness.
- 4.5.22 Negative features (Figure 4.7) were spread relatively evenly across six categories: mismatch; lack of clarity; absences; poor quality writing; poor design; and inappropriate detail or emphasis. *Mismatches* were found between: conclusions

and evidence; design and interpretation; introduction and discussion; material flagged up in abstract and text; and results and interpretation. In addition, there was some evidence of mismatch between the level of detail provided and likely reader skill or interest. For example, too much statistical knowledge was required of readers in one or two instances. *Lack of clarity* involved conceptual confusion, poor explanations of results, poor definition of terms or simply too much detail included. The category *absences* included lack of: critical scrutiny, particularly relating to contentious issues; exemplars e.g. feedback sheets, questionnaires, criteria; research evidence/references; sufficient practical detail to enable replication; and inter-rater reliability testing. In addition, some studies were rated as having poor quality writing, poor design or inappropriate emphasis and detail.

Figure 4.7



Concluding observations

4.5.23 There is much that is encouraging contained within the sub-set of entries dealing with the involvement of students in assessment in the Innovative Assessment database. Many authors are locating their studies within a theoretical framework and/or the literature on the topic. The reasons many authors introduce their innovation concern benefits to students, to teachers and to assessment itself. There is much useful advice and suggestions for further implementations to be found. However, the sub-set also illustrates less useful features. While many entries provide a helpful template for aspiring researchers and practitioners and stimulate thought, others are lacking in these areas. As we have seen, positive and negative features identified in entries by reviewers are roughly balanced. Given that so many examples are categorised as accounts of practice, it may be helpful to authors to provide some guidelines on what constitutes a good account.

4.6 Use of new technology in assessment

Defining and delimiting the theme

- 4.6.1 This theme concerns innovation in assessment that involves the use of new technology, where assessment is mediated through or delineated by the technological medium. Much of the use of new technology in assessment is associated with the increasing use of virtual or managed learning environments in the provision of part or all of the course content.

Items retrieved by type or nature of innovation

- 4.6.2 A total of 46 items were retrieved and reviewed. Out of those, eight were not included as in six cases assessment was not the focus of the article and in the other cases either the article dealt with further not higher education or was concerned with ICT in education in a general context. In considering type of innovation across the 38 articles, 23 were concerned specifically with multiple choice type assessment and its variants. Seven were concerned with computer-mediated communication in the form of discussion boards and possible collaborative work by students, and the ways that could be incorporated into assessments. Three involved the creation of a website as an assessment, and a further three explored the management of institutional risk in the context of electronic assessment. Two articles looked at the submission of portfolios as assessment with some or all of the elements presented electronically, and one article examined the submission of assessment using multimedia on a CD Rom.

Items retrieved by subject area

- 4.6.3 There is a wide variety of subject areas represented by the articles retrieved. Many of them (15) have no subject-specific grouping, but instead have a general, institutional or broad variety of subjects as their focus. Five articles are in the field of Information and Computer Science (with one overlapping with Psychology and one with Education). Education; Bioscience; and Mathematics, Statistics and Operational Research (one overlap with Engineering); are each the focus of three articles. Business Management, Accounting and Finance; Engineering; and Health Science and Practice each feature in two articles. Geography, Earth and Environmental Science; Built Environment; Psychology; and History, Classics and Archaeology each feature in one article.

Items retrieved by genre

- 4.6.4 By genre the articles fall in to the following categories

- account of practice (22) two overlap with empirical study and one with conceptual categories
- enhancement project (6) one overlaps with guidelines.
- conceptual/ theoretical (4)
- empirical study (3)
- evaluation (2)
- guidelines (2) one overlaps with compendium of evolving practices.

- 4.6.5 *Accounts of practice.* Some of these articles used the literature as a basis for the account and to suggest implications for future practice. Emerging themes from the accounts of practice articles read suggest that the use of new technology in assessment:
- opens the dialogue about assessment and learning
 - highlights the importance of feedback and the desirability of swift feedback for a positive effect on learning
 - emphasises the value of aligning assessment with learning outcomes
 - highlights the importance of staff familiarity with the medium and the possible provision of training and support
 - highlights the importance of feedback, support and facilitation for students, including the provision of opportunities to practise assessment tasks
 - emphasises the resource implications and the often 'front loaded' nature of introducing technologically based assessment.
 - highlights the use of computer-mediated communication and collaborative work as a resource base for assessment.
 - the value of ICT in diagnostic testing
 - creative possibilities in assessment such as web page design
 - the need to design assessment to minimise risk.
- 4.6.6 *Enhancement projects.* All five of these articles emphasise the importance of assessment's *conventional wisdom* particularly with respect to learning outcomes; incremental support for learning through assessment; the richness and timeliness of feedback; the importance of allocating marks to engender student motivation; and opening the dialogue about learning.
- 4.6.7 *Conceptual.* Three of these conceptual articles would benefit from a firmer grounding in the literature. Practical implications for assessment per se are difficult to find in two of the articles, with the emphasis in one being on the security issues surrounding on-line assessment and another being a descriptive account of the potential uses of multimedia in assessment.
- 4.6.8 *Empirical study.* Two of these studies were well referenced and suggested how using technology in assessment could be taken forward, one by studying the use of adaptive testing, and the other by exploring the concept of contributions to electronic discussion boards being seen as written communication rather than speech by students. The third article in this genre was disappointingly descriptive.
- 4.6.9 *Evaluation.* One article does not specifically deal with a particular innovative assessment but with the use of the electronic medium in assessment. This is a robust evaluation of the use of a variety of computer aided assessments at Southampton University. It suggests both pragmatic recommendations for future work and implications for good practice. These implications for good practice are congruent with principles that apply in assessment more generally, not just in the

electronic medium e.g. the importance for programme teams to have a common view of the distribution of responsibilities within the team, and the possibilities that systems failure can cause even assessments with the very best design to become ineffective, highlighting the need to build in contingencies for assessment should this happen. The other evaluates the implementation of a variety of uses of electronic assessment and makes sound recommendations for practice.

- 4.6.10 *Guidelines.* One deals specifically with the biosciences and offers a showcase of case studies and specific advice to teachers. The other is a notably high quality overview of CAA in HE and offers practical advice at an institutional level about how to take CAA forward.

Drivers to innovation

- 4.6.11 The drivers of the research have been practically oriented in the majority of the papers. This practical orientation takes the form of either a review of established practice, or establishing a practice with a view to taking the use of new technology in assessment forward in a general or subject-specific sense. Similarly, the drivers to the innovation can be seen to have implications at either an institutional level or for individual practitioners. In some cases, external funding has been the driver to the innovation, for example HEFCE and TLTP funding. Despite the preponderance of articles that use the development of multiple-choice types of assessment as their driver, there are also a number of articles in which technology used in the context of enriching learning through, for example, discussion, is taken forward to see how it can be used within the context of assessment.

Points of reference

- 4.6.12 In many cases the conceptual and empirical points of reference can be identified, including the work around models of computer-based learning, teaching and assessment (Laurillard); constructive alignment (Biggs); social constructivism (Vygotsky, Bruner, Biggs, Lebow); the feedback literature (Black and William, Higgins et al); computer based assessment (Bull, McKenna, Brown, Race); and academic literacies (Lea). There were also however a number of articles which were based on current practice or previous work by the author and showed little acknowledgement of more wide-ranging conceptual or empirical points of reference.

Implications drawn by authors

- 4.6.13 What is notable is that many of the emerging themes gathered from the implications suggested by the authors reflect established good practice in assessment more generally, not just in the electronic environment including:
- alignment of assessment to learning outcomes
 - clarity of purpose of assessment
 - the importance of clear and transparent assessment and marking criteria
 - the need for adequate resource to establish electronic assessment

- acknowledgement that the design and implementation of assessment using the electronic medium is often 'front loaded' and has implications for staff workload and for student support as it is implemented for the first time
- the need for training of both staff and students in the use of electronic assessment and feedback and the management of the ensuing dialogue about learning.

Concluding observations

4.6.14 The quality of the articles varied enormously from accounts of practice, some which were mainly descriptive, others robust with well-evidenced conclusions and recommendations for future practice. The enhancement projects were all conceptually based on literature and utilised a sound methodology. Overall, the articles of all genres offered suggestions and implications for future practice as summarised above. Some of the most striking work was in the area of the institutional implications for using technology in assessment, where models were suggested to minimise risk for the institution, teacher and student (articles by *Zakrzewski et al, 1998, 2000, 2003) and further practical recommendations in the suggested institutional guidelines by *Stephens, Bull and Wade (1998). Work in the field of academic literacies has implications for the use of collaborative and discussion board based assessment, where the use of electronic discussions as a written medium is being explored (*Goodfellow and Lea, 2005). Apart from these examples and some of the work on the use of ICT to allow students to present their work in the form of a website (*France and Ribchester, 2004), it can be seen that although the use of technology may be innovative in assessment, the type of assessment being used may not be so.

4.7 **What guidance and feedback is given, by what means, and when**

Defining and delimiting the theme

4.7.1 This theme concerns innovation in the provision of guidance from tutors about expectations for assessed work, and in the generation and provision of feedback to students (usually in the form of comments and marks, grades or criterion-specific ratings) on the quality of the work they have submitted. Guidance and feedback are generally seen as central to the important *formative* function of assessment, which is crucial in helping students to achieve high-quality learning outcomes. The main areas of overlap of this theme are with student involvement in assessment – in the form of peer feedback – and with the use of new technology, especially with respect to the role of multiple-choice questions (MCQs) in self-testing and formative feedback. It also intersects with two of the themes not singled out in the present review for more in-depth analysis: *modes of assessment and balance between them*, and *criteria used in the assessment and their relative weighting*).

Items retrieved by type or nature of innovation

4.7.2 A total of 37 items were retrieved and judged appropriate to this theme, while a further eight were recategorised or rejected. Across the 37 items recorded, about two-thirds were concerned with the provision of feedback, with varying emphases on or attention to students' perceptions and experiences of feedback, their use of

it, effects on performance and approaches to assignments of enhancing feedback and/or students' grasp of tutors' expectations and assessment criteria, and analyses of tutors' feedback comments. Eight items were concerned with particular strategies for providing, communicating or enhancing feedback, including the use of computer-based spreadsheet feedback, multiple-choice questions, model answers, one-minute papers in lectures, evolving tutorial files and forms of technology-supported feedback. Five items focused on guidance in the context of supervision, either in a work setting or, in the other four instances, in relation to undergraduate dissertations or final-year projects. Two of the articles retrieved were concerned specifically with postgraduate courses, and one of these focused particularly on international students' grasp of academic expectations in a UK setting. One compendium item included two case summaries on the use of personal response systems (PRS) or *clickers* – electronic remote-control devices used in classroom settings to sense and collate students' answers to multiple-choice questions – as a source of feedback in lectures.

Items retrieved by subject area

- 4.7.3 By far the largest subject grouping was that of Business, Management, Accountancy and Finance (10 publications); followed by Bioscience (4) and Health Science and Practice (3). Other subject areas represented were: Physical Sciences (2); Psychology (2); as well as Social Science (1); Economics (1); Music (1); Humanities (1); Maths for Scientists (1); and Sports Science (1). Twelve of the publications were not linked to any particular subject areas or disciplines.

Items retrieved by genre

- 4.7.4 With respect to genre, ten publications were categorised as accounts of practice and 12 as empirical studies, but it was sometimes difficult to draw a hard-and-fast distinction between these two genres. For instance, five of the publications deemed primarily accounts of practice had also presented significant empirical data, and of those categorised as empirical studies, seven had at the same time focused on practice within a quite limited context. Another publication, ostensibly empirically based, was ultimately categorised as a guide to practice rather than an empirical study since it was geared to a teaching rather than research audience, and linked illustrative quotes from interviews with students to an emerging set of guidelines rather than presenting a systematic analysis of the data collected. All told, there were two sets of guidelines and two guides to professional practice, as well as four reviews of the literature, two contributions to theory, four enhancement projects (which of course have some characteristics in common with empirical studies and with accounts of practice) and one compendium of practices. However, two of the four reviews to varying extents also sought to contribute to theoretical advance, while two of them articulated principles to underpin practice, in one case focusing on the provision of feedback and in the other on the more specific issue of approaches to engaging students with feedback.
- 4.7.5 Where data had been collected, the methods deployed were usually student and staff questionnaires or interviews (individual or group), complemented in a small number of instances by analysis of grades or marks. Only four studies examined tutors' written feedback comments. However, in almost every instance of data-gathering (and regardless of whether the item concerned could be characterised as an empirical study or an account of practice) sample sizes tended to be relatively small, and quite circumscribed (e.g. confined to a particular course

setting, subject area, level/year of study and/or institution). Despite these evidential limitations, however, there was little attempt in many of the items retrieved to consider explicitly the robustness of the data, or to suggest what constraints on wider generalisability (e.g. to other subject areas or institutional settings, or even extrapolating across year-groups) would be called for. Nonetheless this was not, it should be stressed, universal: indeed, a minority of studies did demonstrate a much greater caution and a more appropriately reflexive approach. Moreover, in most of the accounts of practice, the inherent problems of the *originator-as-interlocutor* were not even broached. In some (fortunately few) instances, staff evidently felt it appropriate to speak for the student perspective – usually suggesting great student satisfaction with the innovation described – even though tangible indications of a student voice were absent.

Drivers to innovation

- 4.7.6 Where they were made explicit, the drivers to innovation appeared to range widely. While no single stimulus or cluster of stimuli emerged, factors that did recur to some extent were systemic rather than localised: concerns about the formidable challenges of providing feedback of sufficient quantity and quality in a period of constrained resources, especially following the advent of much larger and more diverse student intakes; widespread evidence of dissatisfaction on the part of many students with provision of feedback and guidance; and staff concerns (often reported but seldom substantiated) that some students at least seemed to show little interest in engaging with or even scrutinising feedback comments. Feedback and guidance were also perceived as warranting greater empirical study, whether for systemic reasons (a general lack of research findings) or more particularised ones (e.g. a lack of solid data on whether a guidance and feedback initiative – or a specific method of generating or communicating feedback – had had the impact envisaged). A less frequently mentioned but nonetheless significant driver was 'pull' rather than 'push': a desire to capitalise on opportunities opened up by developments in information and communications technologies.

Points of reference

- 4.7.7 There were few attempts to define or delimit feedback or guidance, while the points of reference for the studies varied a great deal. The conceptual frameworks espoused included the general literature on assessment in higher education (e.g. the work of Knight, Brown [both P. and S.], Gibbs); research and theory (not just in higher education) on formative assessment (especially the work of Sadler, Black and William, Yorke); empirical studies on feedback in higher education (e.g. Hounsell, Lea, Higgins et al.); constructivist (Bruner, Vygotsky, Biggs) and socio-cultural (Gipps) perspectives on learning; participatory theories of learning (e.g. Lave and Wenger on communities of practice, Northedge on disciplinary discourses); the 'academic literacies' approach (e.g. Lea and Street, Lillis, Ivanic); and learning and self-regulation (Pintrich, Butler and Winne).
- 4.7.8 It should also be observed that even though conceptual diversity is the norm rather than the exception in UK educational research, academic development and day-to-day practice, there were few indications (looking across this full corpus of references) of the existence of a canon, i.e. of a core of empirical studies or conceptual standpoints that no-one active in the field and claiming to be reasonably well-read could choose to ignore. Indeed, in some of the items retrieved, it seemed fortuitous what was, or was not, cited as a conceptual or

empirical point of reference. However, it must continually be borne in mind that many of the authors of these publications were subject practitioners rather than educational researchers. Set against this necessary caution, it can be argued, a core if not a canon of key conceptual and empirical studies could be identified by those well-acquainted with the literature on the field, and a challenge for the Academy might therefore be to consider whether and how this core might be made more accessible to subject teachers and academic developers.

Implications drawn by authors

- 4.7.9 These were very varied and often highly localised. Two implications that came across quite strongly, however, were the need for much greater recognition of the inherent challenges of providing constructive and effective feedback, and the desirability of greater attention to tutor-student dialogue and interchange about expectations. Underpinning both these implications was a recognition that the provision of effective guidance and feedback is not a technical matter but embedded within a complex and dynamic process of intersubjective communication within subject communities. The latter was also reflected in the theoretically oriented publications, albeit considered from different conceptual standpoints.

Concluding observations

- 4.7.10 What proved difficult to ignore in analysing this corpus of publications was twofold. First was the extent to which much of the literature – including empirical studies – had been produced by university teachers whose subject area was not education (and who therefore might well have only a limited understanding of educational research methodology and a limited acquaintance with and access to the relevant literature). It would not therefore be appropriate to evaluate these publications as if they had been produced by established educational researchers. Equally, however, it needs to be better-recognised that subject teachers can bring to their reports a *street-credibility* – the *wisdom of experience*, a grasp of what is practicable, a deep acquaintance with their subject and its pedagogy – which the educational researcher is rarely in a position to match.
- 4.7.11 A second concluding observation concerns the pervasiveness of accounts of practice, which can claim to represent a distinctive genre for reasons explored elsewhere in this report, but which is dogged by the problem that the conventions for writing and publishing accounts of practice are neither clear nor widely understood or agreed. We must of course remind ourselves – as noted in the introduction to this report – that the purposes underlying the reporting of accounts of practice are not necessarily those of empirical studies, nor are subject practitioners necessarily seeking to derive from accounts of practice what they – and educational researchers generally – typically seek from empirical studies. Yet without some consensus on conventions and standards of reporting, it is difficult to see how the overall quality of publications within the genre can be raised. Nonetheless, a small number of the accounts of practices reviewed here did appear to offer excellent illustrations of how an account of practice could be well reported. Models and exemplars of this and other genres could therefore be identified.
- 4.7.12 A further comment should be made about the empirical research literature. Not only was it very largely small-scale, as already noted, but there were no examples

of two kinds of studies, the results of which could be considered crucial to advancing understanding within this field: longitudinal research tracking students' evolving perceptions and utilisation of guidance and feedback over the span of a degree programme; and research focusing not only on the how and what of tutors' feedback comments but also on how these are interpreted and addressed by the students.

5 Analysis of selected subjects

5.1 Introduction

5.1.1 This section of the report presents overviews of publications retrieved on seven subject areas chosen so as to be representative of the broader span of disciplines focused upon in this search-and-review. The subject areas chosen were: Bioscience; Business, Management, Accountancy and Finance; Geography, Earth and Environmental Sciences; Health Sciences and Practice; History, Classics and Archaeology; Languages, Linguistics and Area Studies; and Psychology. Reference is also made as appropriate to other relevant subject-specific publications to be found on the website of the Academy Subject Centre concerned. It should be noted that (as with the thematic critiques in section 4 above), since the final phases of work on both the database and the overviews had by necessity to be done concurrently, the numbers of publications discussed under each subject area below will not precisely match those given above in section 3. Concluding observations on the subject overviews are made in the closing section of the report (see 6.14-6.15).

5.1.2 The following sections were designed to be read in conjunction with the extensive EndNote database. For reasons of space, and because the database can be searched by the fields described in section 3 above, the analyses which follow focus on providing an overview of the findings rather than providing details of the relevant examples, which can all be found in the database itself.

5.2 Bioscience

5.2.1 The Innovative Assessment database includes 17 articles drawn from the biosciences subject area, representing 4.7% of all the items included. These articles were categorised under the following topics: student involvement in assessment (7 articles); technology in assessment (3); feedback and guidance on assessment; and portfolios and other unconventional writing assignments (two each); and assessment of groups and collaborations, criteria used in assessment, and innovation in assessment holistically (one each). The paper classified under the holistic category was published by the Student Assessment and Classification Working Group. They aimed to help identify discipline-related inconsistencies in assessment practices, and included biosciences as one of their subjects. They reported that coursework marks in one set of subjects, which included the

biosciences, were as much as two-thirds of one honours class higher than marks from examinations; this discrepancy contrasted with the smaller difference in the English and History subject group. Six of the seven items classified within the student involvement with assessment category were concerned, either as their central focus or as a major subsidiary topic, with the derivation and use of marking criteria. Since five of these articles involved the same author (Paul Orsmond), this emphasis on criteria is more likely to reflect an individual's research interest rather than any discipline-specific issue. Four of the papers dealt with aspects of assessing project work, perhaps reflecting the important part that research projects have traditionally played in bioscience degrees.

- 5.2.2 Eleven (58%) of the items were accounts of practice; hence the dominance of this genre in other subject areas was also found here. There were two articles in each of the categories empirical study, enhancement project and compendium of evolving categories. *Assessment & Evaluation in Higher Education* published 14 of the articles. Two items were books from the Higher Education Academy Centre for Bioscience Enhancing Learning Series, two were articles from other peer-reviewed journals and one was a book chapter.
- 5.2.3 The website of the Academy Subject Centre for Biosciences provides access to a wide variety of material of relevance to assessment in the biosciences, most of which is not included in the Innovative Assessment database. The centre has developed an assessment audit tool to help with module and programme design; this encourages the use of a wide range of assessment types. The centre publishes its own electronic peer-reviewed journal, *Bioscience Education e-Journal*, which has included 16 articles addressing assessment so far. The centre also publishes a newsletter (*Centre for Bioscience Bulletin*), which has to date published at least 11 articles about assessment.

5.3 **Business, Management, Accountancy and Finance**

- 5.3.1 The database includes 38 items specifically concerned with the subject area of Business, Management, Accountancy and Finance. This was by far the best-represented subject area in the database, accounting for 14% of the subject-specific items retrieved and 10.2% of the overall total.
- 5.3.2 Within the Innovative Assessment database, the most frequently occurring topics were the provision of guidance and feedback (nine items); assessment criteria and their weighting (nine items); assessment of groups and teams (6); student involvement in assessment (5); enhancing assessment practices (3); and use of new technology in assessment (3). Five items focused on assessment and student diversity, and of these, one was concerned with access students, and four with international students, although three of the latter were by the same author.
- 5.3.3 The most common genre was that of accounts of practice, of which there were 21 examples, although a further nine items classified as chiefly empirical studies were also in part accounts of practice in a particular course setting and locale. There were 15 empirical studies in all – almost all of which were based on relatively small samples in a particular settings – as well as five reports on enhancement projects (i.e. initiatives in which there was collaboration between subject specialists and

academic developers or educational researchers) and two publications which were conceptual in focus. No other genres were represented.

- 5.3.4 Looking beyond the Innovative Assessment database to the website of the Higher Education Academy Subject Centre for Business, Management, Accountancy and Finance would yield further subject-focused materials, in a downloadable form. There is a folder on *assessment* containing some 20 items, about two-thirds of which are concerned with innovative assessment of various kinds, notably computer-aided assessment; groupwork and peer-assessment; self-assessment; and assessing reflective learning (e.g. via learning journals). There is also the electronic *International Journal of Management Education*, launched in 2002, in which innovative assessment is the subject of three of around 60 articles published to date, although two of these also appear in the *assessment* folder. In both sources, accounts of practice predominate, but interestingly there are also two examples of the Subject Centre's own concise guides to professional practice.

5.4 Geography, Earth and Environmental Sciences

- 5.4.1 Twenty-four papers from the Innovative Assessment database come from the subject area of Geography, Earth and Environmental Sciences (GEES). This represents 6% of all entries. The papers represent each of our six key themes, as well as the topic of innovative assessment holistically. Four papers cover categories not analysed in depth in this study. The topic of portfolios accounted for the greatest number of papers from the GEES area (5), followed by groups and collaboration (4), involving students (4), presentations (2), the use of new technology (1) and feedback (1).
- 5.4.2 The majority of papers for GEES are published in the higher education journal dedicated to this area, *The Journal of Geography in Higher Education* (14). *Assessment & Evaluation in Higher Education* is another major source of information for teachers of Geography, Earth and Environmental Sciences as well as a possible location for publishing research in the area (six papers). Single papers dealing with innovative assessment in the GEES area were also found in *Higher Education*, *Studies in Higher Education* and *Teaching in Higher Education*. One book was also found. There may be other sources not identified due to the sampling strategy used in this study. Assessment in the context of fieldwork is particularly well covered (nine papers). The coverage of issues such as gender differences and feminism (4) suggest that the discipline has a tradition of investigation of innovation in assessment and a mature approach to this endeavour. The topical issue of plagiarism was the focus of one paper.
- 5.4.3 This subject area is also well served by the Higher Education Academy Subject Centre for Geography, Earth and Environmental Sciences (GEES) website, (www.gees.ac.uk). There is easy access to information about the organisation, with details of useful contacts, as well as a link to the Higher Education Academy website and the other 23 Subject Centres. There is a news archive, and events page, as well as a useful Resource Database. This contains learning and teaching resources for those who teach and/or support learning in geography, earth, and environmental sciences higher education. The publication of the GEES Subject Centre, *Planet*, is available on-line. The Resources Database has an extensive section of *good practice case studies*. A search of this found 18 dealing with

assessment, 14 of which dealt with peer-assessment. Peer-assessment was combined with self-assessment in two cases, and with aspects of group work in a further five, with fieldwork or posters in a further two case studies each. Single case studies combined peer assessment with on-line study, problem-based learning, presentations and writing, respectively. Some case studies contributed to more than one category. Self assessment was the subject of a further four case studies which focussed on reading skills, reflection, recording skills, portfolios, work placements and Computer Assisted Assessment (CAA). Again, some studies fell into more than one category. A search for *collaborative assessment* resulted in identification of 12 case studies, none of which was exclusively or directly associated with assessment as such, though some may contain aspects of formative collaboration.

5.4.4 The website provides links to the Geography Discipline Network (GDN), which contains the GDN Guide on Assessment with further case studies. In addition to useful general information relating to teaching and learning, the GDN Guides also contain lists of web pages related to assessment in Geography. The link to the *DeLiberations* site was accessed and information about subject-based education journals for Geography, Earth and Environmental Sciences located. Further links to four journals are provided: *Journal of Geography in Higher Education*, *Journal of Geoscience Education*, *Environmental Education Research* and *Ecology and Society*.

5.5 Health Sciences and Practice

5.5.1 From the total number of articles listed in the Innovative Assessment database, 27 (6.7%) of the 403 database entries were related to health sciences and practice, and that constituted the third largest specific grouping. 18 papers were published since 2000 and those that were published before then tended to be related to assessment for competency or related to work-based learning. Nine of those papers were related to the theme of portfolios and other unconventional writing assignments; seven were in the untitled category (but interestingly all of the papers were related to the allied health professions of physiotherapy occupational therapy, radiography or chiropody); five were related to feedback and guidance on assessment; four were related to technology in assessment; and one was on the subject of student involvement in assessment.

5.5.2 Assessment based upon reflection in the form of portfolios, logs or journals spanned the years from 1997-2007. As might be expected, assessment issues related to nursing (5) featured most within this group especially in relation to the use of portfolios. Other health and social care professions featured to a lesser degree. Journals represented in the subject of Health Sciences and Practice revealed that 13 were published in the journal of *Assessment & Evaluation in Higher Education*. Of the rest, there were three chapters in textbooks, two papers in *Teaching in Higher Education*, four papers in nursing journals (*Nursing Education Today* (2) and *Journal of Advanced Nursing* (2)). There was one paper in *Studies in Higher Education*, three papers in *Active Learning in Higher Education* and one in the *British Journal of Educational Technology*.

5.5.3 A wide range of concerns regarding assessment are noted from the papers related to assessment in work-based learning, the failing student, how to assess

reflection, how to assess portfolios, assessment of competence, and issues related to the reliability and validity of a specific type of timed test. Regarding the theoretical background for innovation in assessment, there was a perceptible shift towards collaboration, altering the balance of power in assessment and strengthening the influence of formative assessment. Constructivism remains the most usually stated theoretical underpinning. Promoting reflective practice is the main rationale used for the use of portfolios. Self-assessment did not feature highly in those papers except in one concerning problem-based learning.

- 5.5.4 Issues concerning innovative assessment need to be contextualised within innovative approaches to learning and teaching, and more material might have been expected concerning e-learning. Four papers on web-based assessment practices outlined the way that assessment could be integrated into collaborative learning processes that foster deep learning, enhance students understanding of the research process and support revision strategies. There was only one paper that critically evaluated the contested aspect of the assessment process within problem-based learning. The predominant finding was of non-alignment between philosophy and assessment with students feeling that group work was undervalued. One paper was related to the role of supervision as a possible aid to interprofessional learning at postgraduate level. Again, given the shift in higher education in health and social care to incorporate interprofessional learning within the undergraduate curriculum, it is perhaps surprising that no paper was found on assessment at that level.
- 5.5.5 Feedback from assessment was a common concern related to timeliness, how meaningful it is to students and whether it is subsequently evaluated. Formative feedback needed to be linked to the summative assessment with more attention to *feedforward* discussions, plus tutors needed to offer more responses to encourage reflectivity about the subject matter. There was a perceived art in sharing the purpose of assessment with learners and in offering a place and scaffolding structure to discuss this. With regard to the ways of exploring innovative assessment, qualitative approaches were chosen by most writers if they wished to systematically explore aspects of feedback. This was typically small-scale and related to evaluation of current practice. Case studies, action research, illuminative evaluation, and evaluations formed the main methods of exploration.
- 5.5.6 In exploring further than the database to the website of the Higher Education Academy Subject Centre for Health Sciences and Practice, a number of workshops had been held in 2006 to help develop students' employability skills through using portfolios. In addition, there were a number of conference opportunities to explore e-portfolios in relation to Personal Development Planning (PDP).

5.6 History, Classics and Archaeology

- 5.6.1 Eleven papers on the database are from the subject area of History, Classics and Archaeology, representing less than 4% of the total entries. Of these, ten focused on History and one on Archaeology. Five of the articles were not solely about History or Archaeology, but were also about innovations in other subject areas. Nine of the papers were articles from a range of journals on teaching and learning in higher education, while one was a refereed conference paper and one a set of

guidelines produced within a higher education institution. Six of the papers were classified as empirical studies, four as accounts of practice and one as guidelines. However, there was some overlap in these classifications, with two of the accounts of practice presenting some data, and one of the empirical studies having the purpose of enhancing practice within an institution.

- 5.6.2 A range of themes were covered in the 11 articles, which were classified on the database under writing assignments, presentations, peer-assessment, group assessment, the use of new technology, feedback, criteria used in assessment, plagiarism, assessment regimes and enhancing assessment practices. The innovations in assessment also varied. They included introducing forms of assessment that were new in a particular context, such as student presentations, free-writing exercises, group projects, on-line tasks, multiple-choice questionnaires to provide feedback, giving generic feedback on exam answers, and a series of workshops for students of different aspects of assessment. Some of the innovations were to find out more about practices and policies in assessing History across an institution or a range of institutions, such as staff and student attitudes to plagiarism, criteria used in marking final-year projects, staff gender differences in marking and giving feedback, how grades are distributed, and differences between exam and coursework marks.
- 5.6.3 The drivers for the innovation were equally varied, including wanting to improve the learning and skills of students on a particular course or in an institution, or to help students with the transition from school to university. Other authors were concerned with the consistency and *fairness* of grades or feedback given to students by staff in one institution or more widely. Another was researching staff and student perceptions of plagiarism in order to inform policy and practice within an institution, while several of the studies were comparing assessment practices and regimes across institutions.
- 5.6.4 Given the variation in the topics and themes covered by the papers, it is unsurprising that a wide range of literature was drawn on by the authors. Work on student learning such as approaches to learning, using experience for learning, active learning, Biggs and Collis' SOLO taxonomy, and the pedagogy of technology, formed the conceptual underpinning for many of the innovations. Previous research on assessment was also widely cited, including specific reference as relevant to grading, plagiarism or the pedagogy of technology. There was also reference to policies on quality assurance, benchmarking and transferable skills.
- 5.6.5 Implications drawn by authors from their research were often for practitioners and would be relevant to staff in a range of disciplines, for example, the importance of making grading criteria clear, being positive when giving feedback, or providing help for staff in implementing new methods of assessment. Some of the articles offered examples of how innovations in assessment had helped their students to develop their learning by, for example, providing early feedback or feedback that helped students evaluate their learning, providing opportunities for students to develop their writing or presentation skills, or helping students to be reflective about or take responsibility for their learning. Other implications could be for institutions, such as developing new policies and practices for dealing with plagiarism. Some of the issues raised had wider implications, for example, for reliability when a range of grades were given for the same essay, or where the distribution of marks varied across subjects or between exams and coursework, or

for changing a culture from one where learning was seen as competitive to a collaborative environment.

- 5.6.6 Besides the articles collated on the database, the Higher Education Academy Subject Centre for History, Classics and Archaeology website contains a number of articles and links on the theme of assessment. It provides a short bibliography to some key texts, a briefing paper on assessing oral presentations, a guide on assessing student learning and some case studies on, for example, multiple-choice testing, group projects, reflective logs and audio-visual gobbets in exams.

5.7 Languages, Linguistics and Area Studies

- 5.7.1 Seven of the 317 Innovative Assessment database entries are concerned with the subject area of Language, Linguistics and Area Studies. This represents only 2.2% of the database. Five of the seven database entries were published in *Assessment & Evaluation in Higher Education*, one is in an educational book and one is in *ReCall*, the Journal of the European Association for Computer Assisted Language Learning. In relation to innovative assessment, Languages, Linguistics and Area Studies is not well represented in the general educational journals. Only 1.1% of database references from education journals are concerned with this subject area. All of those articles are from *Assessment & Evaluation in Higher Education*, whereas four of the five searched education journals had not published any article concerned with Languages, Linguistic and Area Studies relevant for the database.
- 5.7.2 One of the articles was not solely about the subject area of Language, Linguistics and Area Studies, but compared data from different subject areas. However, most of the articles stressed that the findings would also be of interest to other subject areas.
- 5.7.3 Looking at the distribution of languages, three articles referred to Language, Linguistics and Area Studies in general without focus on one particular language, three papers related to French and one article discussed Teaching English to Speakers of Other Languages (TESOL). Only two of the modern languages taught at higher education institutions in the UK are represented.
- 5.7.4 With respect to the genre, the majority of papers, five items, were classified as accounts of practice and two items as empirical studies. There was some overlap between these genres as three accounts of practice presented data which went beyond anecdotal evidence.
- 5.7.5 The main themes were student involvement in assessment (3); criteria used in the assessment and their relative weighting (2); use of new technology in assessment (1); and innovation in assessment generally – assessment regimes (1). Two of the papers on student involvement in assessment were from the same author. The innovations were located at course unit and degree programme level within and across institutions. They included:
- self-assessment of graded work after tutor feedback (two items)
 - self- and peer-assessment to prepare students for their research project abroad

- introduction of web-based facilities in English and in the target language French
- development of assessment criteria for an MSc in TESOL
- assessment and recording of transferable skills
- subject-specific marking behaviour in ten subjects, including French, and its implications.

5.7.6 There were a variety and combination of drivers for these innovations. Some authors wanted to improve students' learning and skills on a particular course, or had to adapt assessment to structural changes in the degree programme. Other innovations were policy driven or compared assessment practices across subjects.

5.7.7 The implications drawn by the authors from their research were mostly applicable across subject areas. However, some were specific to Language, Linguistics and Area Studies. One paper highlighted, for example, students need to be more aware of the mistakes caused by their native language in translations. Another article identified transferable skills which are in particular developed through the learning of a foreign language. The wide range of topics researched is reflected in the implications drawn by the authors, which are relevant for practitioners and at institutional level. For example, the importance of linking self- and peer-assessment to the aims and objectives of a course, or students' preference for self-assessment with integrated tutor feedback to self-assessment without tutor feedback, would be of relevance to practitioners. The distribution of marks across disciplines and its implications for degree classifications, however, raises questions at institutional level.

5.7.8 The website of the Higher Education Academy Subject Centre for Languages, Linguistics and Area Studies provides a number of subject-focused material on innovative assessment. Some *Guides to Good Practice*, papers and project reports focus on innovative assessment practices, and some discuss it among other aspects of teaching and learning. The relevant material is mainly concerned with portfolios including the European Language Portfolio, but self- and peer-assessment, the Statement of Relevance as a reflective tool, assessment issues for dyslexic language learners, and assessment of presentations and tandems are also discussed. In addition, the website gives access to a wide range of teaching materials including online multiple-choice tests, which can be used for diagnostic, formative or summative assessment.

5.8 **Psychology**

5.8.1 The Innovative Assessment database includes ten items specifically concerned with the subject area of Psychology and this represents 3% of the database entries. Of the ten entries half were articles published in the key journal *Assessment & Evaluation in Higher Education*, one was a journal article from *Studies in Higher Education* and one from *Higher Education*. Of the remaining three items, two are from the Centre for Excellence in Teaching and Learning at Northumbria University with Liz McDowell as lead author, and the remaining one is an article from the journal of LTSN Psychology.

- 5.8.2 Four of these psychology-based articles were in research themes which were outside the themes critically analysed for the database, and the remainder included the use of new technology in assessment, student involvement in assessment, feedback and assessment holistically.
- 5.8.3 The most common genre was empirical studies with four out of the ten articles falling into this category. In the database overall the most prevalent genre is accounts of practice, suggesting that empirical studies are more prevalent in this subject area than many others. There were three accounts of practice, with enhancement projects and reviews of the literature each being represented by one article. The final article was a mixture of two categories – a compendium of evolving practice and guidelines.
- 5.8.4 Beyond the database, the Higher Education Academy Subject Centre for Psychology website has a core team based in York with a partner site at Strathclyde University. The Psychology Network homepage links through teaching practice to teaching practice resources, and a section on assessment and examining. This section contains a wide variety of downloadable resources including:
- a range of articles covering different types of assessment within psychology including: different types of assessment from criterion and norm referencing through exams; peer assessment; computer aided assessment and associated feedback software; on-line formative assessment; and essays to the appropriate application of a variety of these assessment strategies. There is also an article on the very specific area of postgraduate Health Psychology
 - an archive of assessment resources relating specifically to Psychology
 - references and links to generic assessment material from a range of projects, CETL materials and the Enhancement Themes in Scotland.
- 5.8.5 In addition to these web-based resources there are a variety of publications from the Psychology Network, the journal of which has featured a special issue which dealt exclusively with assessment within the discipline. It is interesting that both the review of the literature and the publications from the CETLs identified in the database overlap with items in the Academy Subject Centre list of relevant material.

6 **Conclusions, implications and recommendations**

- 6.1 The results of this review indicate clearly that the UK literature on innovative assessment – in common with the literature of education more generally (Hounsell, 1987) – is large and buoyant as well as richly varied and diffuse. Focusing only on the last decade, and notwithstanding the constraints outlined above, it proved possible to identify and record some 317 publications, stretching across (albeit unevenly) the full span of subject areas, focusing on a wide array of themes, representing a diversity of genres and communicative purposes, and

spread across a myriad of publication channels and sources. Before reviewing these features of the innovative assessment literature in greater depth, however, some caveats and cautions are necessary.

- 6.2 First, the review was undertaken within pre-set boundaries of time and other resources. The resulting database and analysis therefore represent what was achievable within those constraints (and without precise foreknowledge of how many items might be retrieved), rather than the outcomes of a fully exhaustive search-and-review of the salient literature. That meant in particular that the core higher education journals were given relatively (though not exclusively) greater priority in the search-and-review strategy, which have contributed in some degree to the predominance of journal articles in the database (accounting for some four-fifths of the overall total). Priority was also given to publication sources that were cross-disciplinary: indeed, the focus has been on literature *across* rather simply *within* the disciplines, i.e. publications that seek to communicate beyond the subject area within which the work reported was conducted. Consequently, the database tends to under-represent subject-specific publications less visible to those outside the subject communities concerned, as well as the 'greyer' literature (especially in the form of conference proceedings which are not publicly indexed). No doubt the full corpus of potentially relevant publications is therefore somewhat larger, though what has not been captured is not necessarily absent from our review, since significant work in the grey literature often finds its way subsequently into more mainstream outlets, just as noteworthy initiatives first reported within a particular subject area can go on to attract a wider audience.
- 6.3 Secondly, it should be reiterated that we have not surveyed the incidence of innovative assessment in UK higher education over the last decade as such, but its documentation in the literature. Indeed, it seems reasonable to assume that initiatives that have been formally recorded in this way make up a minority rather than a majority of the initiatives which have been attempted across British higher education.
- 6.4 Thirdly, as we indicated in the introductory section of this report, innovation in assessment is a slippery notion which is interpreted in ways that are bound up with the particular settings within which it arises, on the one hand, and on the other, are viewed through the eyes of its beholders. We have therefore judged it appropriate to take a relatively inclusive definition of what counts as 'innovative', leaving it open to end-users to make finer-grained distinctions between what might be more – or less – innovative for themselves and the settings within which they work.
- 6.5 Fourthly, it should be borne in mind that there is much non-UK literature on innovative assessment in higher education that can be also of interest and value to a UK audience. Australian publications – e.g. work by Boud (1986, 1995) and Sadler (1989, 1998), and two compendia of changing practices (Nightingale et al., 1996, James et al. 2002) – have been notably influential, and there has been growing attention to publications originating in South East Asia, continental Europe and the US (e.g. Carless et al., 2006; Segers, Dochy and Cascallar, 2003; Wilson and Scalise, 2006). In other words, the potential resource that might be of assistance to the would-be UK innovator is larger than the present database alone; and likewise, the publications in the database have potential value far beyond these shores.

Innovation by theme

- 6.6 Within the classification scheme developed for and refined within the course of the review, 12 main categories were differentiated. A total of eight themes spread across six categories accounted for nearly four-fifths of the total number of relevant publications retrieved and recorded in the database: student involvement in assessment (58 items); portfolios and similar assignments (49 items); the use of new technology in assessment (45 items); the provision of guidance and feedback (40 items); assessment of groups and collaboration (37 items); how assessment tasks are organised, aligned and managed (31 items); criteria and their relative weighting (29 items); and presentations and other non-written assignments (27 items). However, the fact that all but two of these themes (organisation of assessment tasks, criteria and their weighting) had been focused on for fuller analysis may have boosted their relative standing somewhat, since relevant material was more actively searched for.
- 6.7 To seasoned observers of assessment trends, this constellation of the most prominent themes within the literature will come as little surprise. A striking example of a topic that had not been anticipated, however, was consideration of the implications for assessment practices and processes of greater student diversity in addressing the needs of UK students from non-traditional backgrounds, students with disabilities, or students from overseas countries, particularly where English was not their first language.
- 6.8 Other emerging innovative trends can most readily be found within particular themes (see section 4 above) because each encompassed a cluster of varied innovative directions rather than a single routeway. A review of the keywords deployed in the database would also offer a means of pinpointing other new developments.
- 6.9 Finally, it should be noted, the Innovative Assessment themes and descriptors would together provide a valuable baseline against which to track future trends in documented innovations in assessment.

Theme drivers

- 6.10 Looking across the thematic analyses presented in section 4, drivers of or stimuli to innovation in assessment of a wide variety of kinds were observed. Some of these were theme-specific, but many were common to several themes. The chief recurring stimuli included: a desire to recraft assessment practices the better to reflect contemporary mass higher education with its large and diverse cohorts of students and pressures on resources; an interest in enhancing current practices in ways that capitalised upon developments in information technology or in methods of assessment; a heightened emphasis on developing a range of transferable skills that graduates would need in the workplace and for lifelong learning; and a more general commitment to sustaining and enhancing the quality of students' learning, which was sometimes linked to a particular pedagogical philosophy — to promote active, reflective learning, for instance, to foster collaboration rather than competition, or to contribute to democratising assessment.

Thematic points of reference

- 6.11 In the thematic analyses, a similarly rich variety was also evident in relation to the points of reference adduced by the authors. These could draw on *how to* guides to teaching and assessment in higher education, literature from the particular subject area concerned, the authors' previous work on the topic, or accounts of practice by others using the same or a similar approach. Equally, they could cite the conceptual literature on assessment (notably formative assessment), empirical studies of assessment and student learning, or the wider pedagogical literature on theoretical perspectives, including, amongst others, constructivism, self-regulation, identity, and academic literacies. A profusion of the latter would not be surprising given Tight's finding, in a review of the higher education literature, that nearly 90 different theoretical standpoints could be identified in a sample of 104 journal articles where theory had been made explicit (Tight, 2004, p. 407).
- 6.12 Across the themes, however, there were also indications of substantial variation in the degree to which authors were knowledgeable about or familiar with relevant sources, whether conceptual, empirical, methodological or practical. In some instances, references were simply sparse; somewhat more frequently, what had been cited seemed fortuitous or opportunistic, rather than reflecting a judicious selection from key sources. That this should have been the case is not surprising, given that so many of the authors were from a subject area other than education. Nonetheless it does have important consequences, for it works against a truly cumulative literature and the evolution of a widely shared understanding of what is known and understood within the field. The Academy, we suggest, may therefore wish to consider what strategies might help to promote the latter. The present review and database have a valuable contribution to make, but might have greater impact if allied to briefing seminars for prospective author-innovators, subject advisers and academic developers.

Theme implications

- 6.13 The thematic analyses also reviewed the practical, procedural and policy implications identified by the authors of the publications retrieved. As might be expected, these implications differed widely in ways that were closely bound up with the theme concerned and with practices in the particular course settings where the innovations had taken place. However, there were also important implications of wider and more general significance, of which three stood out. These were: the necessity of careful groundwork, briefing and training of both students and staff (the latter especially where course teams were themselves large and diverse) if the initiatives reported were to take firm root; the need to attend as appropriate to other salient aspects of course organisation and management, including time and resources (both of which often had to be front-loaded for innovations such as those which were ICT-linked); and an alertness to implications for change at the departmental, faculty or institutional level that would need to be addressed if some innovations in assessment were to flourish in other courses and programmes of study.

Innovation by subject area

- 6.14 Analysis of documented innovations by subject area showed that the best-represented subject areas were those of Business and Management (38 items), Education (32 items) and Health Sciences and Practice (27 items). There were seven subject areas represented by four or fewer publications, including: Philosophy and Religious Studies; Hospitality, Leisure, Sport and Tourism; Law; Built Environment; Physical Sciences; Economics; and Materials. As these two listings would indicate, there were no obvious patterns of incidence related to relative age of a discipline, vocational versus non-vocational fields, or to broader subject groupings (e.g. humanities and social sciences as compared to science and technology).
- 6.15 How far these patterns of distribution might be considered a proxy measure of the scale or scope of innovation in assessment within a given subject area, is of course open to question, as already noted. However, the degree to which innovative assessment is documented within a subject area or discipline may be indicative of the vigour with which innovative initiatives are reported and debated, while a relative dearth of documented initiatives might act as a brake on potential innovation, by furnishing few exemplars from which other colleagues in the same subject area might learn. If this is indeed the case, the Academy and its Subject Centres might wish to consider the merits and feasibility of a capacity-building strategy in which subject areas with relatively low levels of documented innovation were given assistance in surveying and recording changing practices by cognate subject areas with comparatively high outputs.

Innovation by genre and fitness-for-purpose

- 6.16 A guiding principle of this review has been to take full cognizance of the extent to which the literature of innovative assessment in higher education is characterised, not only by a profusion of origins but also by a diversity of destinations, needs and end-uses. This in turn means eschewing the assumption that the rigour and value of any publication depends on the extent to which it conforms to the conventions for published research and scholarship, i.e. that it has been produced by researchers and principally for researchers, and follows established, research-driven norms and practices.
- 6.17 A key focus of the review has therefore been the genres of the publications surveyed, and taking in a range of genres that are both characteristic of the innovative assessment literature and reflect a diversity of destinations and end-uses that lie in practice and policy as well as research. In consequence, we did find a relatively large number of empirical studies (74 items, making up about one-fifth of the database), reviews of the literature (14 items), and contributions to theory/conceptualisation (51 items) – all traditionally associated with research, but here as likely to be authored by subject practitioners as mainstream educational researchers. But we also found numerous examples of genres that were more overtly practice-generated and practice-focused: guidelines (13 items), guides to professional practice (9 items), commentaries/opinion pieces (4 items), and two *Cinderella* genres which seem particularly associated with the innovative assessment literature: accounts of practice (141 items) and compendia of practices (7 items), as we have termed them. Indeed, as these figures will suggest, accounts of practice proved to be the most common genre in the database, accounting for a remarkable two-fifths of the overall total. There was

also an important emergent genre, the enhancement project (19 items) that brings together the complementary expertise of the subject practitioner and the educational researcher. There were no examples of policy documents, and comparatively few instances of evaluations (7 items), perhaps because assessment initiatives tended to be self-evaluated rather than undertaken by a third-party.

Accounts of practice

- 6.18 As was apparent at various points in the thematic analyses presented in section 4, accounts of practice can have potentially substantial shortcomings, and of two principal kinds. Firstly, they are characteristically self-reported, self-monitored and self-evaluative, which means that their authors are to a considerable extent *parti pris*. It is in their authors' interests to put the initiative in the best possible light, or at least to minimise or make light of its shortcomings; and the reliability of their account depends to a critical extent on what they choose to focus on or emphasise, what data or evidence germane to the effectiveness of the initiative they choose to select and present, and how this is interpreted by them against the backcloth of the course setting concerned. Secondly, as far as we have been able to ascertain, there are no well-understood or readily accessible conventions for reporting accounts of practice, with the consequence that there are no clear rules-of-thumb or guiding principles which authors, referees or editors might rely on. This lack of guidelines may be all the more problematic because the authors of accounts of practice will generally be much more familiar with the particular standards and conventions prevailing in their own discipline or subject area than with those applicable to the higher education literature.
- 6.19 Given these shortcomings, it hardly seems surprising that accounts of practice are often looked down on, or disregarded as significant or *serious* contributions to the literature. Yet they can also have very considerable strengths. They are not simply aspirational or visionary, but practice-embedded reports of authentic course-based experiences, and documented by established subject insiders rather than transient educational researchers. Accounts of practice may thus have a high degree of *street-credibility* for fellow-practitioners. They also play a vital role in documenting and disseminating within and beyond a subject what has been learnt from developments in practices, and in inspiring others to emulate them.
- 6.20 Furthermore, while their general quality could be variable, it was possible in our thematic analyses to identify many good examples of accounts of practice, and so to tease out what main features of accounts of practice seemed likely to inspire confidence in readers and end-users, and thus maximise their potential value in the advancement of practice and policy. Four key features were uppermost in the best accounts of practice:
- a. their authors made explicit the nature of their own involvement in the initiatives being related and were alert to the potential limitations this might raise
 - b. they attempted to locate the account within the relevant literature, by drawing on what was already known and deploying an appropriate conceptual framework

- c. they sought to highlight distinctive features of the course, subject and institutional setting within which the initiative reported was implemented, and to consider its outcomes in the light of these
 - d. they made explicit the nature, scope, sources and robustness of the data reviewed or specifically gathered, in arriving at any conclusions about the effectiveness of the initiative or considering wider implications.
- 6.21 In the light of these observations, and given the prevalence of accounts of practice in the innovative assessment literature, the Academy might wish to consider how their overall quality and their potential value to practitioners could best be enhanced. One possible way forward would be to draw up a code of practice or set of guidelines, ideally in consultation with relevant journal editors and professional and scholarly bodies. This could valuably be complemented by fuller guidance materials, illustrated by reference to exemplary accounts of practice, and/or a rolling programme of seminars and workshops (see Table 6.1). In either case, it might be fruitful to differentiate between the needs of those who wished to generate and contribute to the literature of innovative assessment, and those who wished to learn from that literature through appropriately *artful* and skilled reading. Greenhalgh's (2001) justly renowned guide to informed and critical reading of papers in Medicine would provide an excellent model in developing resources on the range of genres typically found in the pedagogical literature of higher education.

Empirical studies

- 6.22 Empirical studies were also strongly represented in the Innovative Assessment database, as already noted, could be authored by subject teachers as well as academic developers and higher education researchers, and were sometimes hard to distinguish from accounts of practice that had gathered and analysed data from beyond the immediate setting within which the initiative reported had taken place. Sample sizes were for the most part small and restricted in one or more respects, as were the range of course or institutional settings surveyed. Yet while the better examples of empirical enquiry were appropriately alert to, and reflectively guarded about, their limitations, there was a proneness amongst the less satisfactory examples towards incautious generalisation regarding institution, subject area, student characteristics (backgrounds and aspirations, gender, first-language) level of study, staff expertise and experience, or departmental resources, facilities and cultures.
- 6.23 Enhancing this genre might lie in two directions. One would be for the Academy, as with accounts of practice, to promote measures (e.g. resource materials, short courses) to build research capacity amongst subject teachers and academic developers. Another would be for the Academy to try to *scale up* sample sizes, by acting as a *dating agency* or broker to encourage and support alliances between aspiring researchers with similar interests (see Table 6.1).
- 6.24 Neither strategy, however, may suffice to address a fundamental weakness in research into innovative assessment in higher education; and that is the virtual absence of follow-up studies that track the impact of assessment approaches (or indeed teaching-learning activities) on students' progression over successive

years or levels of study. This gap is especially serious in relation to the evolving interchange of and responsibility for feedback, and for the development of skills in communication and teamwork.

- 6.25 Perhaps more controversially, and notwithstanding both the practical and ethical challenges entailed, there would seem to be untapped potential for controlled studies of innovative assessment initiatives, the findings from which could provide a valuable complement to findings from more widely used research approaches.

Enhancement projects and evaluations

- 6.26 A brokerage role akin to that suggested above might also be helpful in relation to enhancement projects and evaluations. In the case of the emergent genre of enhancement projects, there were relatively fewer examples to be found than of empirical studies or accounts of practice, but the overall quality appeared to be much higher than was the case for the latter, suggesting that such partnerships between subject practitioners and colleagues with expertise in higher education R & D could fruitfully be more proactively encouraged and supported (see Table 6.1).

End-uses of the literature

- 6.27 A report of this kind inevitably has to be made – and be guided by – assumptions about the end-uses of the literature of innovative assessment by subject teachers, by academic managers and policy-makers, and not least by those with roles in supporting academic development, whether in institutional teaching-learning or quality enhancement support units, or in the Academy itself and its network of Subject Centres and advisers. At present, however, such assumptions have of necessity to be based chiefly on the accumulated experience of the review team themselves as researchers, academic developers and subject teachers, and from their interactions with colleagues. This is not, it hardly needs saying, a desirable situation. There is a pressing need for systematic empirical investigation of how the pedagogical literature of higher education is accessed, made sense of and put to use for practical and policy-related purposes. Happily, there is a small number of examples from higher education (Kezar, 2000) and other sectors of education (e.g., Nutley, Percy-Smith and Solesbury, 2003; John and Prior, 2003; DETYA, 2001) where this has been attempted and which could assist in exploring possible ways of proceeding.

Review impact and sustainability

- 6.28 Finally, both this review and the database which accompanies it have an important role to play not only in mapping the focus and incidence of documented innovations in assessment in higher education, but also in itself making a contribution to capacity-building. The construction of the database should make it much more straightforward for those pursuing and documenting innovative initiatives to familiarise themselves with what has been attempted elsewhere and what lessons can be drawn, within and across subject areas. At the same time, the analysis of the items retrieved by genre should be of assistance to end-users

of the literature by alerting them to the distinctive communicative and evidential strengths and limitations of publications from different genres.

- 6.29 It is also hoped that the typology of genres developed for this particular review may prove to be applicable and worthwhile in other analyses of the wider pedagogical literature of higher education.

Table 6.1: Recommendations for Action

<i>Suggested action</i>	<i>Taken forward by whom?</i>
1. Building capacity to review and document innovation in assessment in subject areas where there is currently little or no such literature.	SCs (in partnership where appropriate with SDAs and relevant CETLs)
2. Enhancing the quality of accounts of practice (e.g. by drawing up a code of practice, setting guidelines, producing guidance materials, running seminars and workshops).	An ad hoc forum comprising representatives of the Academy, editors of key teaching-learning journals, relevant PSBs and SQE
3. Enhancing the quality of empirical studies (e.g. through award-bearing programmes, resource materials, short courses).	ADCs with well-established research expertise HEA
4. Increasing sample sizes for empirical studies by brokering and supporting alliances between prospective researchers.	SCs HEA PSBs
5. Encouraging the development of enhancement projects (e.g. through partnerships between subject practitioners and colleagues with expertise in higher education R&D).	ADCs CETLs SQE
6. Supporting research on how the pedagogical literature of higher education is accessed, made sense of and put to use for practical and policy-related purposes.	HEA PSBs Research funding bodies

- ADCs Units and centres within universities concerned with academic/educational development and enhancing the quality of learning and teaching
- CETLs Centres for Excellence in Teaching and Learning

HEA	The Higher Education Academy
PSBs	e.g. SRHE, SEDA
SDAs	Subject and discipline-focused societies and associations
SCs	The Higher Education Academy's Subject Centres
SQE	The Scottish Enhancement Themes programme

7 Acknowledgements

We would like to express our thanks to the Higher Education Academy, which commissioned this study and gave it financial support; to Edinburgh, Napier and Glasgow Caledonian Universities, for the significant commitments of staff time, expertise and other resources without which this study would not have been possible; and to the members of our Steering Committee – Professors Liz McDowell, Lin Norton, Brenda Smith and Mantz Yorke – who generously found the time and space to contribute their many helpful reflections, observations and insights.

8 References

- Birenbaum, M. and Dochy, F.J.R.C. (eds.) (1996) *Alternatives in Assessment of Achievements. Learning Processes and Prior Knowledge*. Boston: Kluwer Academic Publishers.
- Boud, D. (1986) *Implementing Student Self-Assessment*. HERDSA Green Guides, no. 5. Kensington NSW: Higher Education Research and Development Society of Australasia.
- Boud, D. (1995) *Enhancing Learning through Self-Assessment*. London: Kogan Page.
- Brown, S. and Glasner, A. (eds.) (1999) *Assessment Matters in Higher Education*. Buckingham: SRHE and Open University Press.
- Bryan, C. and Clegg, K., (eds.) (2006) *Innovative Assessment in Higher Education*. London: Routledge.
- Carless, D., Joughin, G., Liu, N.-F. et al., (eds.) (2006) *How Assessment Supports Learning: Learning-Orientated Assessment in Action*. Hong Kong: Hong Kong University Press. http://www.hkupress.org/asp/bookinfo.asp?PD_NUM=9622098231
- Cotterill, S. Bradley, P., and Hammond, G. (2006) 'ePortfolios supporting assessment in complex educational environments' (Chapter 17, 191-199) In: Bryan, C. and Clegg, K. *Innovative Assessment in Higher Education*. Routledge: London.
- DETYA (2001) *Educational Research: In Whose Interests?* (Higher Education Series, report no. 39). Canberra: Department of Education, Training and Youth Affairs, Higher Education Division.
- Ecclestone, K. and Swann, J. (1999) Litigation and learning: tensions in improving university lecturers' assessment practice. *Assessment in Education*. 6(3), 377-389.
- Falchikov, N. (2005) *Improving Assessment through Student Involvement*. London: Routledge-Falmer.

- Gibbs, G. (1995) *Assessing Student Centred Courses*. Oxford: Oxford Centre for Staff Development.
- Greenhalgh, T. (2001) *How to Read a Paper. The basics of evidence-based medicine*. London: BMJ Books.
- Hannan, A. and Silver, H. (2000) *Innovating in Higher Education: Teaching, Learning and Institutional Cultures*. Buckingham: SRHE and Open University Press.
- Hounsell, D. (1987) 'The Supply and Demand for Information about Education'. In: Boyd-Barrett, V. and Braham, P., (eds.) *Media, Knowledge and Power*. London: Croom Helm/Open University, 434-443.
- Hounsell, D. and McCulloch, M. (1999) 'Assessing Skills in Scottish Higher Education'. In: Dunne, E. (ed.) *The Learning Society: International Perspectives on Core Skills in Higher Education*. London: Kogan Page, 149-158.
- Hounsell, D., McCulloch, M. and Scott, M. (1996) (eds.) *The ASSHE Inventory: Changing Assessment Practices in Scottish Higher Education*. Edinburgh and Sheffield: University of Edinburgh, TLA Centre/Napier University/CVCP UCoSDA. Republished electronically in 2002 by the Learning and Teaching Support Network URL: <http://www.ltsn.ac.uk/genericcentre/asshe/>
- Hounsell, D., Xu, R. and Tai, C.M. (2007) *Blending Assignments and Assessments for High-Quality Learning*. (Scottish Enhancement Themes: Guides to Integrative Assessment, no.3). Gloucester: Quality Assurance Agency for Higher Education. ISBN 978 1 84482 649 0.
- James, R., McInnis, C. and Devlin, M. (2002) *Assessing Learning in Australian Universities*. Melbourne: University of Melbourne, Centre for the Study of Higher Education for the Australian Universities Teaching Committee. <http://www.cshe.unimelb.edu.au/assessinglearning>
- John, P. and Prior, J. (2003) 'Conceptions, contentions and connections: how teachers read and understand different genres of educational research.' In: Sutherland, R., Claxton, G. and Pollard, A. *Learning and Teaching Where Worldviews Meet*. Stoke on Trent: Trentham Books, 231-243.
- Kezar, A. (2000) 'Understanding the research-to-practice gap: a national study of researchers' and practitioners' perspectives.' In: Kezar, A. and Eckel, P., (eds.) *Moving Beyond the Gap between Research and Practice in Higher Education*. (New Directions for Higher Education, no. 110). San Francisco: Jossey-Bass, 9-19.
- MacLellan, E. (2004) How convincing is alternative assessment for use in higher education? *Assessment & Evaluation in Higher Education*. 29.3, pp. 311-321.
- McKenzie, J., Alexander, S., Harper, C. and Anderson, S. (2005) *Dissemination, Adoption and Adaptation of Project Innovations in Higher Education*. A report for the Carrick Institute for Learning and Teaching in Higher Education. Sydney: University of Technology Sydney. <http://www.carrickinstitute.edu.au/carrick/go/home/grants/pid/344>
- Nightingale, P., Te Wiata, I., Toohey, S., Ryan, G., Hughes, C., and Magin, D. (1996) *Assessing Learning in Universities*. Sydney: University of New South Wales Press.
- Nutley, S. M., Percy-Smith, J. and Solesbury, W. (2003) *Models of research impact: a cross-sector review of literature and practice*. London: Learning and Skills Research Centre.
- Nutley, S., Walter, I. and Davies, H.T.O. (2003) From knowing to doing: a framework for understanding the evidence-into-practice agenda. *Evaluation*, 9(2), 125-148.
- Sadler, D.R. (1989) Formative assessment and the design of instructional systems, *Instructional Science*. 18, 119-144.

- Sadler, D.R. (1998) Formative assessment: revisiting the territory. *Assessment in Education*. 5 (1) 77-84.
- Segers, M., Dochy, F. and Cascallar, E., (eds.) (2003) *Optimising New Methods of Assessment: In Search of Qualities and Standards*. Dordrecht: Kluwer.
- Southwell, D., Gannaway, D., Orrell, J., Chalmers, D. and Abraham, C. (2005) *Strategies for Effective Dissemination of Project Outcomes*. A report for the Carrick Institute for Learning and Teaching in Higher Education. Brisbane: University of Queensland, Flinders University and Queensland University of Technology.
<http://www.carrickinstitute.edu.au/carrick/go/home/grants/pid/344>
- Struyven, K., Dochy, F. and Janssens, S. (2003) 'Students' perceptions about new modes of assessment in higher education: a review'. In: Segers, M., Dochy, F. and Cascallar, E., (eds.) *Optimising New Methods of Assessment: In Search of Qualities and Standards*. Dordrecht: Kluwer. 171-223.
- Swales, J. M. (1990) *Genre Analysis. English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Tight, M. (2003) *Researching Higher Education*. Maidenhead: SRHE and Open University Press/McGraw-Hill.
- Tight, M. (2004) Research into higher education: an a-theoretical community of practice? *Higher Education Research and Development*, 23(4), 395-411.
- Wilson, M. and Scalise, K. (2006) Assessment to improve learning in higher education: the BEAR assessment system. *Higher Education*. 52(4), 635-663.